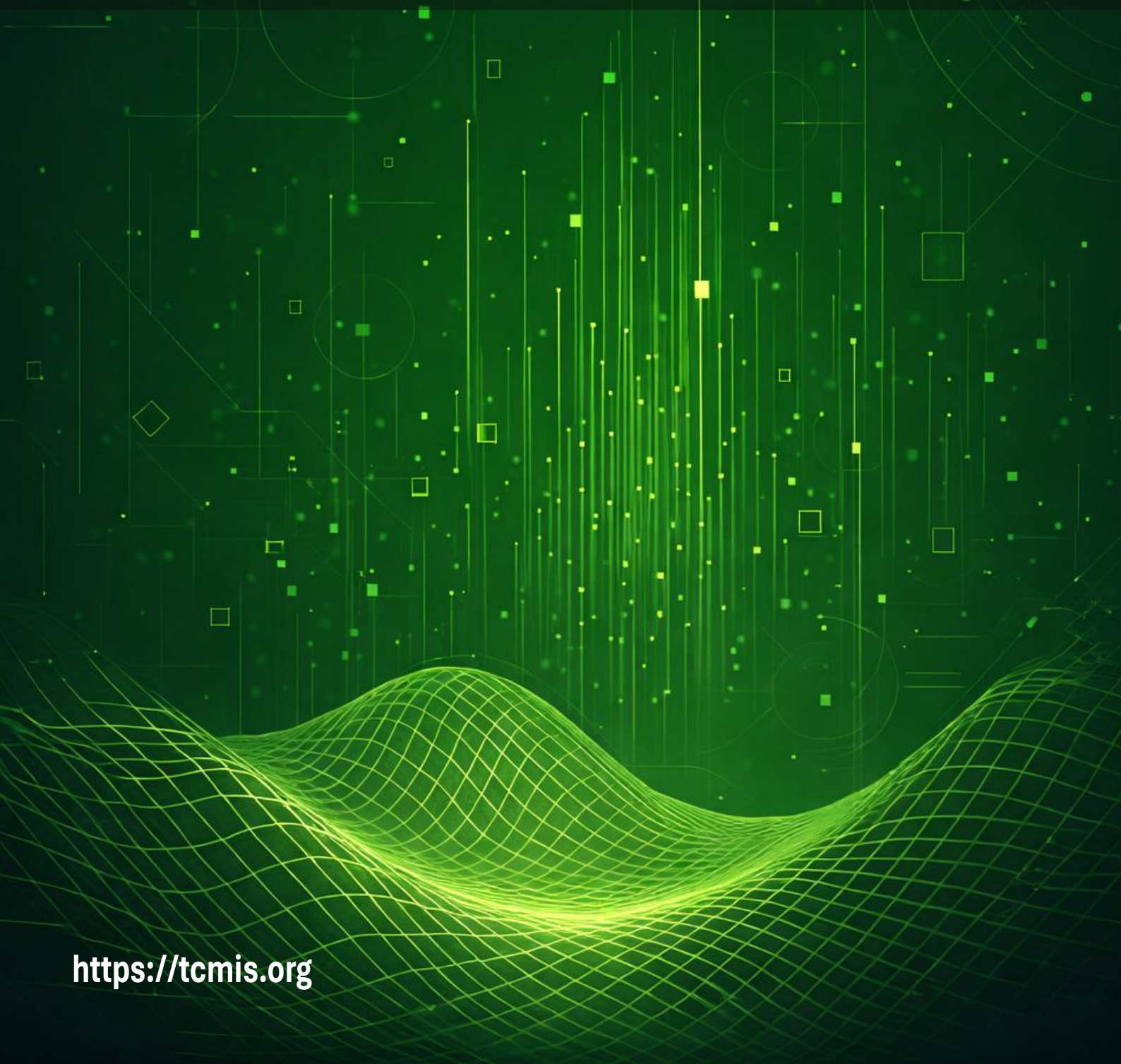
eISSN: 3108-5202

TRANSACTIONS ON

COMPUTATIONAL MODELING AND INTELLIGENT SYSTEMS

Volume 2, January 2026



https://tcmis.org

Transactions on Computational Modeling and Intelligent Systems

Volume 2, January 2026

eISSN:

3108-5202

Editor-in-Chief:

Prof. Dumitru Baleanu
Institute of Space Sciences, Bucharest, Romania

Official Title:

Transactions on Computational Modeling and Intelligent Systems

Official Abbreviation:

Trans. Comput. Model. Intell. Syst.

Publisher:

Gülşen Yaman

Publication Period:

Quarterly (in January, April, July, October)

Webpage:

https://tcmis.org

Email:

editor@tcmis.org

Area Editors

Operational Research and Optimization

Gerhard W. Weber, Poznan University of Technology, Poland

Applied Mathematics and Control

Necati Ozdemir, Balikesir University, Turkiye
Zakia Hammouch, Moulay Ismail University, Morocco

Computational Modelling & Simulation

Guo-Cheng Wu, Chongqing University of Posts and Telecom., China

Intelligent Systems

Mukund N. Janardhanan, University of Warwick, United Kingdom

Editorial Board *(in alphabetical order)*

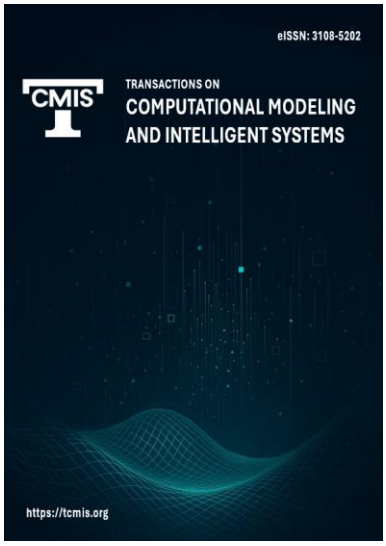
Amar Debbouche, Guelma University, Algeria
Zakia Hammouch, Moulay Ismail University, Morocco
Hossein Jafari, Department of Mathematics, University of South Africa, South Africa
Amin Jajarmi, University of Bojnord, Iran
Mukund N. Janardhanan, University of Warwick, United Kingdom
Yeliz Karaca, University of Massachusetts, United States
Luis Vázquez Martínez, Universidad Complutense de Madrid, Spain
Necati Ozdemir, Balikesir University, Turkiye
Mirco Peron, NEOMA Business School, France
Anett Rácz, University of Debrecen, Hungary
Jocelyn Sabatier, Bordeaux Université, France
Achyut Shankar, Bennett University, India
Gerhard W. Weber, Poznan University of Technology, Poland
Guo-Cheng Wu, Chongqing University of Posts and Telecommunications, China

English Language Editor *(in alphabetical order)*

Dilek Inan, Izmir Democracy University, Turkiye

Abstracted & Indexed in:

Google Scholar, Harvard Hollis, OpenAIRE, Research Commons, Worldcat OCLC



Volume 2, January 2026

Transactions on Computational Modeling and Intelligent Systems

eISSN: 3108-5202

Webpage: https://tcmis.org

An international peer-reviewed open-access journal published quarterly

CONTENTS (Article ID, Title and Authors)

Research Articles

10051 Multi-composite general neural network approximation over finite dimensional Banach spaces

George A. Anastassiou

10043 Advanced modeling for fire risk evaluation and intelligent preventive system

Sabina Umirzakova, Shakhzod Javliev, Adkhambek Madaminov, Omadjon Urishev, Dilshoda Kurbonalieva, Ayhan Istanbullu, Akmalbek Abdusalomov

10012 Uncovering the significance of undone animals' flesh and early treatment on the dynamics of ascariasis using mathematical analysis

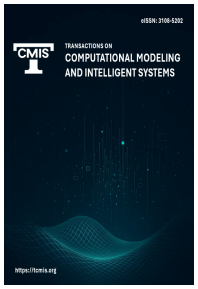
James Andrawus, Ibrahim Ibrahim Muhammad, Felix Yakubu Eguda, Abdullahi Yusuf

10028 A selective review of modern stochastic modeling: SDE/SPDE numerics, data-driven identification, and generative methods with applications in biomathematics

Yassine Sabbar, Kottakkaran Sooppy Nisar

10024 Enhancing pharmacokinetic modeling with the Upadhyaya transform and machine learning: A synergistic approach

Prabakaran Raghavendran, Yamini Parthiban, Dinesh Thakur, Jothivelu Thiravidarani



Transactions on
Computational Modeling
and Intelligent Systems

Volume 2, January 2026, Article 10051
https://doi.org/10.65112.tcmis.10051

eISSN: 3108-5202
www.tcmis.org

RESEARCH ARTICLE

Multi-composite general neural network approximation over finite dimensional Banach spaces

George A. Anastassiou*

Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, U.S.A.
**Corresponding Author. Email: ganastss@memphis.edu (G. A. Anastassiou)*

Article Information

Received: 27 December 2025
Accepted: 10 January 2026
Published 12 January 2026

AMS Classification:
41A17, 41A25, 41A99, 46B25

Abstract

The functions under approximation here have as a domain a finite dimensional Banach space with dimension $N \in \mathbb{N}$ and are with values in $\mathbb{R}^N$. Exploiting some topological properties of the above we are able to perform multi-composite general Neural Network multivariate approximation to the above functions. The treatment is quantitative. We produce multivariate multi-composite general Jackson type inequalities involving the modulus of continuity of the function under approximation. The established convergences are pointwise and uniform. Our technique is expected to lead to accelerated speeds of convergence.

Keywords: Finite dimensional Banach spaces, multi-composite general neural network operators approximation, modulus of continuity, multi-composite general accelerated approximation

1. Introduction

Multi-composition of activation functions in neural networks involves applying different activation functions across layers or even within the same layer to introduce varied non-linearities, allowing models to learn complex patterns better, often seen in advanced architectures like: Long Short-Term Memory (LSTM) networks or through techniques like "Activation Ensembles" to boost performance by blending functions (e.g., Sigmoid, Tanh, ReLU) for improved learning and tackling issues like vanishing gradients, leading to stronger, more versatile models than a single function might provide. The benefits follow:

Enhanced Non-linearity: Different functions capture different types of non-linear relationships, allowing the network to model more intricate data.

Improved Performance: Ensembles of different activation functions often outperform networks using just one, especially on complex tasks or smaller datasets.

Specialized Functions: Combining standard functions (like ReLU for speed) with others (like Sigmoid/Tanh for range control) offers flexibility.

Applications:

Recurrent Networks: LSTM networks inherently use multiple functions (sigmoid, tanh) to manage cell state and gates.

Advanced Ensembles: Methods like "Activation Ensembles" learn weights for multiple functions at each neuron, creating a richer activation.

Hybrid Layers: Concatenating outputs from different activation functions within a layer or using them across different parts of a deep network.

As a conclusion: While a single, well-chosen function (like ReLU) often suffices, using multiple activations, either sequentially or combined, provides a powerful way to boost a neural network's representational power and accuracy, moving beyond simpler "vanilla" network designs.

The author in [1] and [2], see Chapters 2-5 was the first to establish neural network approximation to continuous functions with rates by very specifically defined neural network operators of Cardaliaguet-Euvrard and "Squashing" types, by employing the modulus of continuity of the engaged function or its high order derivative, and producing very tight Jackson type inequalities. He treats there both the univariate and multivariate cases. The defining these operators "bell-shaped" and "squashing" functions are assumed to be of compact support.

Again the author inspired by [3], continued his studies on neural network approximation by introducing and using the proper quasi-interpolation operators of sigmoidal and hyperbolic tangent type which resulted into [4], by treating both the univariate and multivariate cases.

The author continued this trend of approximation and published the extensive monographs [5–9], studying in depth and as wide as possible the neural network approximations. Our activation function here is a multi-composite general activation function.

The multivariate operators we use are: the normalized, the quasi-interpolation, the Kantorovich and the quadrature types. Our multi-composite technique is expected to accelerate immensely the speed of convergence. Our neural networks here are from a Banach space of dimension N to $\mathbb{R}^N$, as they are homeomorphic, and this is the main key to this work.

So all results here are via multivariate multi-composite Jackson type inequalities, studying quantitatively the rate of convergence to the unit operator.

Here we are dealing with one hidden layer feed-forward multi-composite neural networks.

A multilayer feed-forward neural network can be defined as follows (with $m \in \mathbb{N}$ hidden layers):

Let $x \in \mathbb{R}^s$; $s \in \mathbb{N}$, where $x = (x_1, ..., x_s)$; $\alpha_j, c_j \in \mathbb{R}^s$; $b_j \in \mathbb{R}$, with $0 \leq j \leq n$, $n \in \mathbb{N}$.

Here $\langle \alpha_j \cdot x \rangle$ is the inner product, thus $\sigma(\langle \alpha_j \cdot x \rangle + b_j) \in \mathbb{R}$; and $N_n(x) \in \mathbb{R}^s$, by $c_j \in \mathbb{R}^s$, as it is coming from $N_n(x) = \sum_{j=0}^{n} c_j \sigma(\langle \alpha_j \cdot x \rangle + b_j)$.

We define:

$$N_n^{(2)}(x) = \sum_{j=0}^{n} c_j \sigma(\langle \alpha_j \cdot N_n(x) \rangle + b_j) =$$

$$\sum_{j=0}^{n} c_j \sigma \left(\left\langle \alpha_j \cdot \left(\sum_{j=0}^{n} c_j \sigma(\langle \alpha_j \cdot x \rangle + b_j) \right) \right\rangle + b_j \right).$$

Furthermore, we can define

$$N_n^{(3)}(x) = \sum_{j=0}^{n} c_j \sigma \left(\left\langle \alpha_j \cdot N_n^{(2)}(x) \right\rangle + b_j \right).$$

And, in general we define:

$$N_n^{(m)}(x) = \sum_{j=0}^{n} c_j \sigma \left(\left\langle \alpha_j \cdot N_n^{(m-1)}(x) \right\rangle + b_j \right), \quad \text{for } m \in \mathbb{N}.$$

For more studies in neural networks read [10–14].

2. Background

We need:

Definition 1. *Let $i = 1, 2, ...$, and $h_i : \mathbb{R} \to [-1, 1]$ be general sigmoid activation functions, such that they are strictly increasing, $h_i(0) = 0$, $h_i(-x) = -h_i(x)$, $x \in \mathbb{R}$, $h_i(+\infty) = 1$,*

$h_i(-\infty) = -1$. *Also* h_i *is strictly convex over* $(-\infty, 0]$ *and strictly concave over* $[0, +\infty)$, *with* $h_i^{(2)} \in C(\mathbb{R})$. *Examples* $\tanh x$, $erf(x)$, *etc.*

Clearly here it is $1 \geq h_i(1) > 0$, $i = 1, 2, ...$ Consider $\widetilde{h}_i(x) := \frac{h_i(x)}{h_i(1)}$; $i = 1, 2, ..., x \in \mathbb{R}$. Then $\widetilde{h}_i : \mathbb{R} \to \left[-\frac{1}{h_i(1)}, \frac{1}{h_i(1)}\right]$, $i = 1, 2, ...$, is a sigmoid activation function such that is strictly increasing, $\widetilde{h}_i(0) = 0$, $\widetilde{h}_i(-x) = -\widetilde{h}_i(x)$, $x \in \mathbb{R}$, $\widetilde{h}_i(+\infty) = \frac{1}{h_i(1)}$, $\widetilde{h}_i(-\infty) = -\frac{1}{h_i(1)}$. Also $\widetilde{h}_i$ is strictly convex over $(-\infty, 0]$ and strictly concave over $[0, +\infty)$, with $\widetilde{h}_i^{(2)} \in C(\mathbb{R})$.

Obviously the composition $\widetilde{h}_1|_{[-1,1]} \circ h_2$ is strictly increasing from $\mathbb{R}$ into $[-1, 1]$ and $\left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(0) = 0$. More precisely we notice that $\left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(+\infty) = \left(\widetilde{h}_1|_{[-1,1]}(h_2(+\infty))\right) = \left(\widetilde{h}_1|_{[-1,1]}(1)\right) = \frac{h_1(1)}{h_1(1)} = 1$, so that $\left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(+\infty) = 1$.

And $\left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(-\infty) = \left(\widetilde{h}_1|_{[-1,1]}(h_2(-\infty))\right) = \left(\widetilde{h}_1|_{[-1,1]}(-1)\right) = \frac{h_1(-1)}{h_1(1)} = \frac{-h_1(1)}{h_1(1)} = -1$, so that $\left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(-\infty) = -1$.

Also it holds $\left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(-x) = \left(\widetilde{h}_1|_{[-1,1]}(h_2(-x))\right) = \left(\widetilde{h}_1|_{[-1,1]}(-h_2(x))\right) = -\left(\widetilde{h}_1|_{[-1,1]}(h_2(x))\right)$ $= -\left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(x)$.

That is $\left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(-x) = -\left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(x)$, $x \in \mathbb{R}$.

Also it is clear that $\left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)^{(2)} \in C(\mathbb{R})$.

Next acting over $(-\infty, 0]$: let $\lambda, \mu \geq 0 : \lambda + \mu = 1$. Then by convexity of h_2 there we have

$$h_2(\lambda x + \mu y) \leq \lambda h_2(x) + \mu h_2(y), \ \forall\, x, y \in (-\infty, 0];$$

and

$$\widetilde{h}_1|_{[-1,1]}(h_2(\lambda x + \mu y)) \leq \widetilde{h}_1|_{[-1,1]}(\lambda h_2(x) + \mu h_2(y)) \leq \tag{1}$$
$$\lambda \widetilde{h}_1|_{[-1,1]}(h_2(x)) + \mu \widetilde{h}_1|_{[-1,1]}(h_2(y)) =$$
$$\lambda \left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(x) + \mu \left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(y), \quad \forall x, \ y \in (-\infty, 0].$$

So that $\widetilde{h}_1|_{[-1,1]} \circ h_2$ is convex over $(-\infty, 0]$.

Now we work on $[0, +\infty)$: let $\lambda, \mu \geq 0 : \lambda + \mu = 1$. Then by concavity of h_2 there we have

$$h_2(\lambda x + \mu y) \geq \lambda h_2(x) + \mu h_2(y), \ \forall\, x, y \in [0, +\infty);$$

and

$$\widetilde{h}_1|_{[-1,1]}(h_2(\lambda x + \mu y)) \geq \widetilde{h}_1|_{[-1,1]}(\lambda h_2(x) + \mu h_2(y)) \geq \tag{2}$$
$$\lambda \widetilde{h}_1|_{[-1,1]}(h_2(x)) + \mu \widetilde{h}_1|_{[-1,1]}(h_2(y)) =$$
$$\lambda \left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(x) + \mu \left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)(y), \quad \forall x, \ y \in [0, +\infty).$$

So that $\widetilde{h}_1|_{[-1,1]} \circ h_2$ is concave over $[0, +\infty)$.

Therefore $\left(\widetilde{h}_1|_{[-1,1]} \circ h_2\right)$ is a general sigmoid activation function fulfilling all of the properties we stated in Definition 1.

Similarly, $\left(\widetilde{h}_2|_{[-1,1]} \circ h_3\right) : \mathbb{R} \to [-1, 1]$, fulfills Definition 1, and $\widetilde{h}_1|_{[-1,1]} \circ \left(\widetilde{h}_2|_{[-1,1]} \circ h_3\right)$ does the same.

So, we can write the last one as $\widetilde{h}_1|_{[-1,1]} \circ \widetilde{h}_2|_{[-1,1]} \circ h_3$, and this fulfills Definition 1.

Clearly, inductively we derive that

$$H_\lambda := \widetilde{h}_1|_{[-1,1]} \circ \widetilde{h}_2|_{[-1,1]} \circ \widetilde{h}_3|_{[-1,1]} \circ ... \circ \widetilde{h}_{\lambda-1}|_{[-1,1]} \circ h_\lambda; \ \ \lambda \in \mathbb{N}, \tag{3}$$

fulfills Definition 1, that is a multicomposite general sigmoid activation function from $\mathbb{R} \to [-1, 1]$.

We can rewrite

$$H_\lambda = \frac{h_1}{h_1(1)}|_{[-1,1]} \circ \frac{h_2}{h_2(1)}|_{[-1,1]} \circ \frac{h_3}{h_3(1)}|_{[-1,1]} \circ ... \circ \frac{h_{\lambda-1}}{h_{\lambda-1}(1)}|_{[-1,1]} \circ h_\lambda, \quad \lambda \in \mathbb{N}. \tag{4}$$

Next we consider the function

$$\Phi_\lambda(x) := \frac{1}{4}\left(H_\lambda(x+1) - H_\lambda(x-1)\right) > 0, \ \ \forall\, x \in \mathbb{R}, \ \lambda \in \mathbb{N}. \tag{5}$$

We observe that

$$\begin{gathered} \Phi_\lambda(-x) = \frac{1}{4}\left(H_\lambda(-x+1) - H_\lambda(-x-1)\right) = \\ \frac{1}{4}\left(H_\lambda(-(x-1)) - H_\lambda(-(x+1))\right) = \frac{1}{4}\left(-H_\lambda(x-1) + H_\lambda(x+1)\right) = \\ \frac{1}{4}\left(H_\lambda(x+1) - H_\lambda(x-1)\right) = \Phi_\lambda(x). \end{gathered} \tag{6}$$

That is Φ_λ is an even function,

$$\Phi_\lambda(-x) = \Phi_\lambda(x), \quad \forall\, x \in \mathbb{R}, \ \lambda \in \mathbb{N}. \tag{7}$$

We see that

$$\Phi_\lambda(0) = \frac{H_\lambda(1)}{2}. \tag{8}$$

Let $x > 1$, we have that

$$\Phi_\lambda'(x) = \frac{1}{4}\left(H_\lambda{}'(x+1) - H_\lambda{}'(x-1)\right) < 0,$$

by H_λ' being strictly decreasing over $[0, +\infty)$.

Let now $0 < x < 1$, then $1 - x > 0$ and $0 < 1 - x < 1 + x$. It holds $H_\lambda'(x-1) = H_\lambda'(1-x) > H_\lambda'(x+1)$, so that again $\Phi_\lambda'(x) < 0$. Consequently Φ_λ is stritly decreasing on $(0, +\infty)$.

Clearly, Φ_λ is strictly increasing on $(-\infty, 0)$, and $\Phi_\lambda'(0) = 0$.

Observe that

$$\lim_{x \to +\infty} \Phi_\lambda(x) = \frac{1}{4}\left(H_\lambda(+\infty) - H_\lambda(+\infty)\right) = 0, \tag{9}$$

and

$$\lim_{x \to -\infty} \Phi_\lambda(x) = \frac{1}{4}\left(H_\lambda(-\infty) - H_\lambda(-\infty)\right) = 0. \tag{10}$$

That is the x-axis is the horizontal asymptote on Φ_λ.

As a result Φ_λ is a bell shaped symmetric function with maximum

$$\Phi_\lambda(0) = \frac{H_\lambda(1)}{2}. \tag{11}$$

We need

Theorem 1. *It holds*

$$\sum_{i=-\infty}^{\infty} \Phi_\lambda(x - i) = 1, \ \ \forall\, x \in \mathbb{R}. \tag{12}$$

Proof. As similar to [6], p. 286 is omitted. □

Theorem 2. *We have that*

$$\int_{-\infty}^{\infty} \Phi_\lambda(x)\, dx = 1. \tag{13}$$

Proof. As similar to [6], p. 287, it is omitted. □

So that $\Phi_\lambda(x)$ can serve as a density function in general.

We need

Theorem 3. *([15]) Let $0 < \alpha < 1$, and $n \in \mathbb{N}$ with $n^{1-\alpha} > 2$. Then*

$$\sum_{\substack{k=-\infty \\ : |nx-k| \geq n^{1-\alpha}}}^{\infty} \Phi_\lambda (nx-k) < \left(1 - H_\lambda \left(n^{1-\alpha} - 2\right)\right), \tag{14}$$

and

$$\lim_{n \to +\infty} \left(1 - H_\lambda \left(n^{1-\alpha} - 2\right)\right) = 0. \tag{15}$$

Denote by $\lfloor \cdot \rfloor$ the integral part of the number and by $\lceil \cdot \rceil$ the ceiling of the number.
We also need

Theorem 4. *Let $x \in [a,b] \subset \mathbb{R}$ and $n \in \mathbb{N}$ so that $\lceil na \rceil \leq \lfloor nb \rfloor$. It holds*

$$\frac{1}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \Phi_\lambda (nx-k)} < \frac{1}{\Phi_\lambda (1)}, \quad \forall\, x \in [a,b]. \tag{16}$$

Proof. As similar to [6], p. 289 is omitted. □

Remark 1. *We have that*

$$\lim_{n \to \infty} \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \Phi_\lambda (nx-k) \neq 1, \tag{17}$$

for at least some $x \in [a,b]$.
See [6], p. 290, same reasoning.

Note 1. *For large enough n we always obtain $\lceil na \rceil \leq \lfloor nb \rfloor$. Also $a \leq \frac{k}{n} \leq b$, iff $\lceil na \rceil \leq k \leq \lfloor nb \rfloor$. In general it holds (by (12))*

$$\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \Phi_\lambda (nx-k) \leq 1. \tag{18}$$

We make

Definition 2. *We introduce*

$$\widetilde{Z}(x_1, ..., x_N) := \widetilde{Z}(x) := \prod_{i=1}^{N} \Phi_\lambda (x_i), \quad x = (x_1, ..., x_N) \in \mathbb{R}^N, \ N \in \mathbb{N}. \tag{19}$$

It has the properties:
(i) $\widetilde{Z}(x) > 0, \ \forall\, x \in \mathbb{R}^N$,
(ii)

$$\sum_{k=-\infty}^{\infty} \widetilde{Z}(x-k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} \widetilde{Z}(x_1-k_1, ..., x_N-k_N) = 1, \tag{20}$$

where $k := (k_1, ..., k_n) \in \mathbb{Z}^N$, $\forall\, x \in \mathbb{R}^N$,
hence
(iii)

$$\sum_{k=-\infty}^{\infty} \widetilde{Z}(nx-k) = 1, \tag{21}$$

$\forall\, x \in \mathbb{R}^N$, $n \in \mathbb{N}$,
and
(iv)

$$\int_{\mathbb{R}^N} \widetilde{Z}(x)\, dx = 1, \tag{22}$$

that is $\widetilde{Z}$ is a multivariate density function.

Here denote $\|x\|_\infty := \max\{|x_1|, ..., |x_N|\}$, $x \in \mathbb{R}^N$, *also set* $\infty := (\infty, ..., \infty)$, $-\infty := (-\infty, ..., -\infty)$, *upon the multivariate context, and*

$$\lceil na \rceil := (\lceil na_1 \rceil, ..., \lceil na_N \rceil), \quad \lfloor nb \rfloor := (\lfloor nb_1 \rfloor, ..., \lfloor nb_N \rfloor), \tag{23}$$

where $a := (a_1, ..., a_N)$, $b := (b_1, ..., b_N)$.

We obviously see that

$$\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \widetilde{Z}(nx-k) = \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left(\prod_{i=1}^{N} \Phi_\lambda (nx_i - k_i) \right) =$$

$$\sum_{k_1=\lceil na_1 \rceil}^{\lfloor nb_1 \rfloor} ... \sum_{k_N=\lceil na_N \rceil}^{\lfloor nb_N \rfloor} \left(\prod_{i=1}^{N} \Phi_\lambda (nx_i - k_i) \right) = \prod_{i=1}^{N} \left(\sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor} \Phi_\lambda (nx_i - k_i) \right). \tag{24}$$

for $0 < \beta < 1$ *and* $n \in \mathbb{N}$, *a fixed* $x \in \mathbb{R}^N$, *we have that*

$$\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \widetilde{Z}(nx-k) =$$

$$\sum_{\substack{k=\lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_\infty \le \frac{1}{n^\beta}}}^{\lfloor nb \rfloor} \widetilde{Z}(nx-k) + \sum_{\substack{k=\lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta}}}^{\lfloor nb \rfloor} \widetilde{Z}(nx-k). \tag{25}$$

In the last two sums the counting is over disjoint vector sets of k*'s, because the condition* $\left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta}$ *implies that there exists at least one* $\left| \frac{k_r}{n} - x_r \right| > \frac{1}{n^\beta}$, *where* $r \in \{1, ..., N\}$.

(v) As in [5], pp. 379-380, we derive that

$$\sum_{\substack{k=\lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta}}}^{\lfloor nb \rfloor} \widetilde{Z}(nx-k) \overset{(14)}{<} 1 - H_\lambda \left(n^{1-\beta} - 2 \right), \quad 0 < \beta < 1, \tag{26}$$

with $n \in \mathbb{N} : n^{1-\beta} > 2$, $x \in \prod_{i=1}^{N} [a_i, b_i]$.

(vi) By Theorem 4 we get that

$$0 < \frac{1}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \widetilde{Z}(nx-k)} < \frac{1}{(\Phi_\lambda(1))^N}, \tag{27}$$

$\forall\ x \in \left(\prod_{i=1}^{N} [a_i, b_i] \right)$, $n \in \mathbb{N}$.

It is also clear that

(vii)

$$\sum_{\substack{k=-\infty \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta}}}^{\infty} \widetilde{Z}(nx-k) < 1 - H_\lambda \left(n^{1-\beta} - 2 \right), \tag{28}$$

$0 < \beta < 1$, $n \in \mathbb{N} : n^{1-\beta} > 2$, $x \in \mathbb{R}^N$.

Furthermore it holds

$$\lim_{n \to \infty} \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \widetilde{Z}(nx-k) \neq 1, \tag{29}$$

for at least some $x \in \left(\prod_{i=1}^{N} [a_i, b_i]\right)$.

We need

Definition 3. *Here* $\left(X, \|\cdot\|_\gamma\right)$ *is a Banach space.*

Let $f \in C\left(\prod_{i=1}^{N} [a_i, b_i], X\right)$, $x = (x_1, ..., x_N) \in \prod_{i=1}^{N} [a_i, b_i]$, $n \in \mathbb{N}$ *such that* $\lceil na_i \rceil \leq \lfloor nb_i \rfloor$, $i = 1, ..., N$.

We introduce and define the following multivariate linear normalized neural network operator ($x := (x_1, ..., x_N) \in \prod_{i=1}^{N} [a_i, b_i]$):

$$\widetilde{A}_n(f, x_1, ..., x_N) := \widetilde{A}_n(f, x) := \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} f\left(\frac{k}{n}\right) \widetilde{Z}(nx - k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \widetilde{Z}(nx - k)} =$$

$$\frac{\sum_{k_1=\lceil na_1 \rceil}^{\lfloor nb_1 \rfloor} \sum_{k_2=\lceil na_2 \rceil}^{\lfloor nb_2 \rfloor} \cdots \sum_{k_N=\lceil na_N \rceil}^{\lfloor nb_N \rfloor} f\left(\frac{k_1}{n}, ..., \frac{k_N}{n}\right) \left(\prod_{i=1}^{N} \Phi_\lambda(nx_i - k_i)\right)}{\prod_{i=1}^{N} \left(\sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor} \Phi_\lambda(nx_i - k_i)\right)}. \tag{30}$$

For large enough $n \in \mathbb{N}$ *we always obtain* $\lceil na_i \rceil \leq \lfloor nb_i \rfloor$, $i = 1, ..., N$. *Also* $a_i \leq \frac{k_i}{n} \leq b_i$, *iff* $\lceil na_i \rceil \leq k_i \leq \lfloor nb_i \rfloor$, $i = 1, ..., N$.

When $g_1 \in C\left(\prod_{i=1}^{N} [a_i, b_i]\right)$ *we define the companion operator*

$$\widetilde{A}_n^*(g_1, x) := \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} g_1\left(\frac{k}{n}\right) \widetilde{Z}(nx - k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \widetilde{Z}(nx - k)}. \tag{31}$$

Clearly $\widetilde{A}_n^*$ *is a positive linear operator. We have that*

$$\widetilde{A}_n^*(1, x) = 1, \ \forall\ x \in \prod_{i=1}^{N} [a_i, b_i].$$

Notice that $\widetilde{A}_n(f) \in C\left(\prod_{i=1}^{N} [a_i, b_i], X\right)$ *and* $\widetilde{A}_n^*(g_1) \in C\left(\prod_{i=1}^{N} [a_i, b_i]\right)$.

Furthermore it holds

$$\left\|\widetilde{A}_n(f, x)\right\|_\gamma \leq \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\|f\left(\frac{k}{n}\right)\right\|_\gamma \widetilde{Z}(nx - k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \widetilde{Z}(nx - k)} = \widetilde{A}_n^*\left(\|f\|_\gamma, x\right) \tag{32}$$

$\forall\ x \in \prod_{i=1}^{N} [a_i, b_i]$.

Clearly $\|f\|_\gamma \in C\left(\prod_{i=1}^{N} [a_i, b_i]\right)$.

So, we have that

$$\left\|\widetilde{A}_n(f, x)\right\|_\gamma \leq \widetilde{A}_n^*\left(\|f\|_\gamma, x\right), \tag{33}$$

$\forall\ x \in \prod_{i=1}^{N} [a_i, b_i]$, $\forall\ n \in \mathbb{N}$, $\forall\ f \in C\left(\prod_{i=1}^{N} [a_i, b_i], X\right)$.

Let $c \in X$ *and* $g_1 \in C\left(\prod_{i=1}^{N} [a_i, b_i]\right)$, *then* $cg_1 \in C\left(\prod_{i=1}^{N} [a_i, b_i], X\right)$.

Furthermore it holds

$$\widetilde{A}_n^*(cg_1, x) = c\widetilde{A}_n^*(g_1, x), \ \forall\ x \in \prod_{i=1}^{N} [a_i, b_i]. \tag{34}$$

Since $\widetilde{A}_n^*(1) = 1$, *we get that*

$$\widetilde{A}_n(c) = c, \ \forall\ c \in X. \tag{35}$$

We call $\widetilde{A}_n^*$ *the companion operator of* $\widetilde{A}_n$.

For convenience we call

$$\widetilde{A}_n^{**}(f,x) := \sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} f\left(\frac{k}{n}\right)\widetilde{Z}(nx-k) =$$

$$\sum_{k_1=\lceil na_1\rceil}^{\lfloor nb_1\rfloor}\sum_{k_2=\lceil na_2\rceil}^{\lfloor nb_2\rfloor}\cdots\sum_{k_N=\lceil na_N\rceil}^{\lfloor nb_N\rfloor} f\left(\frac{k_1}{n},...,\frac{k_N}{n}\right)\left(\prod_{i=1}^{N}\Phi_\lambda(nx_i-k_i)\right). \tag{36}$$

$\forall\ x\in\prod_{i=1}^{N}[a_i,b_i]$.
That is

$$\widetilde{A}_n(f,x) := \frac{\widetilde{A}_n^{**}(f,x)}{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\widetilde{Z}(nx-k)}, \tag{37}$$

$\forall\ x\in\prod_{i=1}^{N}[a_i,b_i]$, $n\in\mathbb{N}$.
Hence

$$\widetilde{A}_n(f,x)-f(x) = \frac{\widetilde{A}_n^{**}(f,x)-f(x)\left(\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\widetilde{Z}(nx-k)\right)}{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\widetilde{Z}(nx-k)}. \tag{38}$$

Consequently we derive

$$\left\|\widetilde{A}_n(f,x)-f(x)\right\|_\gamma \overset{(38)}{\leq} \frac{1}{(\Phi_\lambda(1))^N}\left\|\widetilde{A}_n^{**}(f,x)-f(x)\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\widetilde{Z}(nx-k)\right\|_\gamma, \tag{39}$$

$\forall\ x\in\prod_{i=1}^{N}[a_i,b_i]$.

We need

Definition 4. *([6], p. 274) Let M be a convex and compact subset of $\left(\mathbb{R}^N,\|\cdot\|_p\right)$, $p\in[1,\infty]$, and $\left(X,\|\cdot\|_\gamma\right)$ be a Banach space. Let $f\in C(M,X)$. We define the first modulus of continuity of f as*

$$\omega_1(f,\delta) := \sup_{\substack{x,y\in M:\\ \|x-y\|_p\leq\delta}}\|f(x)-f(y)\|,\quad 0<\delta\leq diam(M). \tag{40}$$

If $\delta > diam(M)$, then

$$\omega_1(f,\delta) = \omega_1(f,diam(M)). \tag{41}$$

Notice $\omega_1(f,\delta)$ is increasing in $\delta>0$. For $f\in C_B(M,X)$ (continuous and bounded functions) $\omega_1(f,\delta)$ is defined similarly.

Lemma 1. *([6], p. 274) We have $\omega_1(f,\delta)\to 0$ as $\delta\downarrow 0$ iff $f\in C(M,X)$, where M is a convex compact subset of $\left(\mathbb{R}^N,\|\cdot\|_p\right)$, $p\in[1,\infty]$.*

Clearly we have also: $f\in C_U\left(\mathbb{R}^N,X\right)$ (uniformly continuous functions), iff $\omega_1(f,\delta)\to 0$ as $\delta\downarrow 0$, where ω_1 is defined similarly to (40). The space $C_B\left(\mathbb{R}^N,X\right)$ denotes the continuous and bounded functions on $\mathbb{R}^N$.In this study we cover only the case of $p=\infty$. Here $\|\cdot\|_\infty$ denotes the supremum norm.

Definition 5. *When $f\in C_B\left(\mathbb{R}^N,X\right)$ or $f\in C_U\left(\mathbb{R}^N,X\right)$, we define*

$$\widetilde{B}_n(f,x) := \widetilde{B}_n(f,x_1,...,x_N) := \sum_{k=-\infty}^{\infty} f\left(\frac{k}{n}\right)\widetilde{Z}(nx-k) :=$$

$$\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\cdots\sum_{k_N=-\infty}^{\infty} f\left(\frac{k_1}{n},...,\frac{k_N}{n}\right)\left(\prod_{i=1}^{N}\Phi_\lambda(nx_i-k_i)\right), \tag{42}$$

$n\in\mathbb{N}$, $\forall\ x\in\mathbb{R}^N$, $N\in\mathbb{N}$, the multivariate quasi-interpolation neural network operator.

Also for $f \in C_B\left(\mathbb{R}^N, X\right)$ or $f \in C_U\left(\mathbb{R}^N, X\right)$, we define the multivariate Kantorovich type neural network operator

$$\widetilde{C}_n(f,x) := \widetilde{C}_n(f,x_1,...,x_N) := \sum_{k=-\infty}^{\infty}\left(n^N \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t)\,dt\right)\widetilde{Z}(nx-k) =$$

$$\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}...\sum_{k_N=-\infty}^{\infty}\left(n^N \int_{\frac{k_1}{n}}^{\frac{k_1+1}{n}}...\int_{\frac{k_N}{n}}^{\frac{k_N+1}{n}} f(t_1,...,t_N)\,dt_1...dt_N\right)$$

$$\left(\prod_{i=1}^{N}\Phi_\lambda(nx_i-k_i)\right), \tag{43}$$

$n \in \mathbb{N}$, $\forall\, x \in \mathbb{R}^N$.

Again for $f \in C_B\left(\mathbb{R}^N, X\right)$ or $f \in C_U\left(\mathbb{R}^N, X\right)$, $N \in \mathbb{N}$, we define the multivariate neural network operator of quadrature type $\widetilde{D}_n(f,x)$, $n \in \mathbb{N}$, as follows:

Let $\theta := (\theta_1,...,\theta_N) \in \mathbb{N}^N$, $r = (r_1,...,r_N) \in \mathbb{Z}_+^n$, $w_r = w_{r_1,r_2,...,r_N} \geq 0$, such that

$$\sum_{r=0}^{\theta} w_r = \sum_{r_1=0}^{\theta_1}\sum_{r_2=0}^{\theta_2}...\sum_{r_N=0}^{\theta_N} w_{r_1,r_2,...,r_N} = 1;$$

$k \in \mathbb{Z}^N$, and

$$\delta_{nk}(f) := \delta_{n,k_1,...,k_N}(f) := \sum_{r=0}^{\theta} w_r f\left(\frac{k}{n}+\frac{r}{n\theta}\right) =$$

$$\sum_{r_1=0}^{\theta_1}\sum_{r_2=0}^{\theta_2}...\sum_{r_N=0}^{\theta_N} w_{r_1,r_2,...,r_N} f\left(\frac{k_1}{n}+\frac{r_1}{n\theta_1},...,\frac{k_n}{n}+\frac{r_N}{n\theta_N}\right), \tag{44}$$

where $\frac{r}{\theta} := \left(\frac{r_1}{\theta_1},...,\frac{r_N}{\theta_N}\right)$.

We set

$$\widetilde{D}_n(f,x) := \widetilde{D}_n(f,x_1,...,x_N) := \sum_{k=-\infty}^{\infty}\delta_{nk}(f)\,\widetilde{Z}(nx-k) = \tag{45}$$

$$\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}...\sum_{k_N=-\infty}^{\infty}\delta_{n,k_1,...,k_N}(f)\left(\prod_{i=1}^{N}\Phi_\lambda(nx_i-k_i)\right),$$

$\forall\, x \in \mathbb{R}^N$.

In this work we treat the case of $\left(X, \|\cdot\|_\gamma\right) = \left(\mathbb{R}^N, \|\cdot\|_2\right)$, $N \in \mathbb{N}$; where $\|\cdot\|_2$ is the Euclidean norm.

Next, we describe the main ideas of this work.

Remark 2. *Let $(E, \|\cdot\|_1)$ be a Banach space of finite dimension N, i.e. $\dim E = N$, $N \in \mathbb{N}$.*

It is known that any two Banach spaces of the same finite dimension are linearly homeomorphic. Hence $(E, \|\cdot\|_1)$ is linearly homeomorphic to $\left(\mathbb{R}^N, \|\cdot\|_2\right)$: so let $\psi : (E, \|\cdot\|_1) \to \left(\mathbb{R}^N, \|\cdot\|_2\right)$ be the linear homeomorphism.

So let $y \in E$, then $\psi(y) = x \in \mathbb{R}^N$, and back to $y = \psi^{-1}(x) \in E$, that is $\psi^{-1} : \mathbb{R}^N \to E$ is $(1-1)$ and onto map, and both ψ, ψ^{-1} are continuous maps.

Let now $f : E \to \mathbb{R}^N$ be a continuous map, then $f \circ \psi^{-1}$ is a continuous map from $\mathbb{R}^N$ into $\mathbb{R}^N$. We call $g(x) := f\left(\psi^{-1}(x)\right)$, $\forall\, x \in \mathbb{R}^N$.

If $\|f(y)\|_2 \leq M$, $\forall\, y \in E$, so is $\left\|f\left(\psi^{-1}(x)\right)\right\|_2 \leq M$, $\forall\, x \in \mathbb{R}^N$, where $M > 0$.

That is $\left\|\|f(y)\|_2\right\|_\infty < \infty$, implies $\left\|\left\|f \circ \psi^{-1}\right\|_2\right\|_\infty < \infty$.

In case f is a uniformly continuous map, since (see below) ψ, ψ^{-1} are also uniformly continuous, we obtain that $g = f \circ \psi^{-1}$ is a uniformly continuous map.

If $\prod_{i=1}^{N}[a_i,b_i]\subset\mathbb{R}^N$, *we can consider* $g|_{\prod_{i=1}^{N}[a_i,b_i]}$, *which is continuous on a bounded set, that is a uniformly continuous map.*

It is known that every linear function on a finite dimensional normed vector space is uniformly continuous, based on the fact that all norms there are equivalent. Thus, ψ,ψ^{-1} *are uniformly continuous.*

Definition 6. *Let* $f:E\to\mathbb{R}^N$ *be continuous.*

We define the following neural network linear operators ${}_i\widetilde{L}_n(f):E\to\mathbb{R}^N$, $i=1,2,3,4$; $y\in E$, *as follows:*

I)

$$
{}_1\widetilde{L}_n(f)(y)={}_1\widetilde{L}_n(f)\left(\psi^{-1}(x)\right):=\widetilde{A}_n\left(f\circ\psi^{-1},x\right)=
$$
$$
\frac{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left(f\circ\psi^{-1}\right)\left(\frac{k}{n}\right)\widetilde{Z}(nx-k)}{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\widetilde{Z}(nx-k)}= \tag{46}
$$
$$
\frac{\sum_{k_1=\lceil na_1\rceil}^{\lfloor nb_1\rfloor}\cdots\sum_{k_N=\lceil na_N\rceil}^{\lfloor nb_N\rfloor}\left(f\circ\psi^{-1}\right)\left(\frac{k_1}{n},...,\frac{k_N}{n}\right)\left(\prod_{i=1}^{N}\Phi_\lambda(nx_i-k_i)\right)}{\prod_{i=1}^{N}\left(\sum_{k_i=\lceil na_i\rceil}^{\lfloor nb_i\rfloor}\Phi_\lambda(nx_i-k_i)\right)},
$$

$\forall\ x:=(x_1,...,x_N)\in\prod_{i=1}^{N}[a_i,b_i]\subset\mathbb{R}^N$, $\forall\ y\in\psi^{-1}\left(\prod_{i=1}^{N}[a_i,b_i]\right)\subset E$, $n\in\mathbb{N}$.

II) When $f\in C_B\left(E,\mathbb{R}^N\right)$ *or* $f\in C_U\left(E,\mathbb{R}^N\right)$ *we define:*

$$
{}_2\widetilde{L}_n(f)(y)={}_2\widetilde{L}_n(f)\left(\psi^{-1}(x)\right):=\widetilde{B}_n\left(f\circ\psi^{-1},x\right)=
$$
$$
\sum_{k=-\infty}^{\infty}\left(f\circ\psi^{-1}\right)\left(\frac{k}{n}\right)\widetilde{Z}(nx-k)= \tag{47}
$$
$$
\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\cdots\sum_{k_N=-\infty}^{\infty}\left(f\circ\psi^{-1}\right)\left(\frac{k_1}{n},...,\frac{k_N}{n}\right)
$$
$$
\left(\prod_{i=1}^{N}\Phi_\lambda(nx_i-k_i)\right),
$$

$\forall\ x\in\mathbb{R}^N$, $\forall\ y\in E$, $n\in\mathbb{N}$.

III) When $f\in C_B\left(E,\mathbb{R}^N\right)$ *or* $f\in C_U\left(E,\mathbb{R}^N\right)$ *we define:*

$$
{}_3\widetilde{L}_n(f)(y)={}_3\widetilde{L}_n(f)\left(\psi^{-1}(x)\right):=\widetilde{C}_n\left(f\circ\psi^{-1},x\right)=
$$
$$
\sum_{k=-\infty}^{\infty}\left(n^N\int_{\frac{k}{n}}^{\frac{k+1}{n}}\left(f\circ\psi^{-1}\right)(t)\,dt\right)\widetilde{Z}(nx-k)=
$$
$$
\sum_{k_1=-\infty}^{\infty}\cdots\sum_{k_N=-\infty}^{\infty}\left(n^N\int_{\frac{k_1}{n}}^{\frac{k_1+1}{n}}\cdots\int_{\frac{k_N}{n}}^{\frac{k_N+1}{n}}\left(f\circ\psi^{-1}\right)(t_1,...,t_N)\,dt_1...dt_N\right)
$$
$$
\left(\prod_{i=1}^{N}\Phi_\lambda(nx_i-k_i)\right), \tag{48}
$$

$\forall\ x\in\mathbb{R}^N$, $\forall\ y\in E$, $n\in\mathbb{N}$.

IV) Again, when $f\in C_B\left(E,\mathbb{R}^N\right)$ *or* $f\in C_U\left(E,\mathbb{R}^N\right)$ *we define:*

$$
{}_4\widetilde{L}_n(f)(y)={}_4\widetilde{L}_n(f)\left(\psi^{-1}(x)\right):=\widetilde{D}_n\left(f\circ\psi^{-1},x\right)=
$$
$$
\sum_{k=-\infty}^{\infty}\delta_{nk}\left(f\circ\psi^{-1}\right)\widetilde{Z}(nx-k)=
$$

$$\sum_{k_1=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} \delta_{n,k_1,...,k_N} \left(f \circ \psi^{-1}\right) \left(\prod_{i=1}^{N} \Phi_\lambda \left(nx_i - k_i\right)\right) \tag{49}$$

$\forall\ x \in \mathbb{R}^N$, $\forall\ y \in E$, $n \in \mathbb{N}$.

We want to study quantitatively the multivariate approximation of ${}_i\widetilde{L}_n\left(f\right)\left(y\right) \to f\left(y\right)$, as $n \to \infty$, for $i = 1, 2, 3, 4$; $y \in E$.

3. Main results

We present our first approximation result.

Theorem 5. *Let $(E, \|\cdot\|_1)$ be a Banach space such that $\dim E = N \in \mathbb{N}$, and ψ be the corresponding homeomorphism from E onto $\left(\mathbb{R}^N, \|\cdot\|_2\right)$. Here $f \in C\left(E, \mathbb{R}^N\right)$, ${}_1\widetilde{L}_n\left(f\right)$ is as in (46), $x \in \prod_{i=1}^{N} [a_i, b_i] \subset \mathbb{R}^N$, and $y \in \psi^{-1}\left(\prod_{i=1}^{N} [a_i, b_i]\right) \subset E$; $0 < \beta < 1$, $n \in \mathbb{N}$ with $n^{1-\beta} > 2$. Then*

(1)

$$\left\|{}_1\widetilde{L}_n\left(f\right)\left(y\right) - f\left(y\right)\right\|_2 \leq$$

$$\frac{1}{\left(\Phi_\lambda\left(1\right)\right)^N}\left[\omega_1\left(f \circ \psi^{-1}, \frac{1}{n^\beta}\right) + 2\left\|\left\|f \circ \psi^{-1}\right\|_2\right\|_\infty \left(1 - H_\lambda\left(n^{1-\beta} - 2\right)\right)\right] =: \sigma_1\left(n\right), \tag{50}$$

and

(2)

$$\left\|\left\|{}_1\widetilde{L}_n\left(f\right) - f\right\|_2\right\|_\infty \leq \sigma_1\left(n\right). \tag{51}$$

We see that $\lim_{n\to\infty} {}_1\widetilde{L}_n\left(f\right) \overset{\|\cdot\|_2}{=} f$, pointwise and uniformly.

Above ω_1 is with respect to $p = \infty$.

Proof. We have that

$$\left\|{}_1\widetilde{L}_n\left(f\right)\left(y\right) - f\left(y\right)\right\|_2 =$$

$$\left\|{}_1\widetilde{L}_n\left(f\right)\left(\psi^{-1}\left(x\right)\right) - f\left(\psi^{-1}\left(x\right)\right)\right\|_2 =$$

$$\left\|\widetilde{A}_n\left(f \circ \psi^{-1}, x\right) - f\left(\psi^{-1}\left(x\right)\right)\right\|_2 = \tag{52}$$

$$\left\|\frac{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left(f \circ \psi^{-1}\right)\left(\frac{k}{n}\right)\widetilde{Z}\left(nx-k\right)}{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\widetilde{Z}\left(nx-k\right)} - \frac{f\left(\psi^{-1}\left(x\right)\right)\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\widetilde{Z}\left(nx-k\right)}{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\widetilde{Z}\left(nx-k\right)}\right\|_2$$

$$\overset{\text{(by (16), (27))}}{\leq} \frac{1}{\left(\Phi_\lambda\left(1\right)\right)^N}$$

$$\left\|\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left(f \circ \psi^{-1}\right)\left(\frac{k}{n}\right)\widetilde{Z}\left(nx-k\right) - f\left(\psi^{-1}\left(x\right)\right)\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\widetilde{Z}\left(nx-k\right)\right\|_2 =$$

$$\frac{1}{\left(\Phi_\lambda\left(1\right)\right)^N}\left\|\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left[\left(f \circ \psi^{-1}\right)\left(\frac{k}{n}\right) - \left(f \circ \psi^{-1}\right)\left(x\right)\right]\widetilde{Z}\left(nx-k\right)\right\|_2 \leq$$

$$\frac{1}{\left(\Phi_\lambda\left(1\right)\right)^N}\left[\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left\|\left(f \circ \psi^{-1}\right)\left(\frac{k}{n}\right) - \left(f \circ \psi^{-1}\right)\left(x\right)\right\|_2 \widetilde{Z}\left(nx-k\right)\right] =$$

$$\frac{1}{\left(\Phi_\lambda(1)\right)^N}\left[\sum_{\substack{k=\lceil na\rceil \\ :\left\|\frac{k}{n}-x\right\|_\infty\le\frac{1}{n^\beta}}}^{\lfloor nb\rfloor}\left\|\left(f\circ\psi^{-1}\right)\left(\frac{k}{n}\right)-\left(f\circ\psi^{-1}\right)(x)\right\|_2\widetilde{Z}(nx-k)\right.$$

$$\left.+\sum_{\substack{k=\lceil na\rceil \\ :\left\|\frac{k}{n}-x\right\|_\infty>\frac{1}{n^\beta}}}^{\lfloor nb\rfloor}\left\|\left(f\circ\psi^{-1}\right)\left(\frac{k}{n}\right)-\left(f\circ\psi^{-1}\right)(x)\right\|_2\widetilde{Z}(nx-k)\right]\overset{(21)}{\le}$$

$$\frac{1}{\left(\Phi_\lambda(1)\right)^N}\left[\omega_1\left(f\circ\psi^{-1},\frac{1}{n^\beta}\right)+2\left\|\left\|f\circ\psi^{-1}\right\|_2\right\|_\infty\left(\sum_{\substack{k=\lceil na\rceil \\ :\left\|\frac{k}{n}-x\right\|_\infty>\frac{1}{n^\beta}}}^{\lfloor nb\rfloor}\widetilde{Z}(nx-k)\right)\right]\overset{(26)}{\le}$$

$$\frac{1}{\left(\Phi_\lambda(1)\right)^N}\left[\omega_1\left(f\circ\psi^{-1},\frac{1}{n^\beta}\right)+2\left\|\left\|f\circ\psi^{-1}\right\|_2\right\|\left(1-H_\lambda\left(n^{1-\beta}-2\right)\right)\right]. \tag{53}$$

□

It follows the next result all over E.

Theorem 6. *Let $(E,\|\cdot\|_1)$ be a Banach space with $\dim E=N\in\mathbb{N}$, and ψ be the corresponding homeomorphism from E onto $\left(\mathbb{R}^N,\|\cdot\|_2\right)$. Here $f\in C_B\left(E,\mathbb{R}^N\right)$, ${}_2\widetilde{L}_n(f)$ is as in (47), $x\in\mathbb{R}^N$, and $y\in E$; $0<\beta<1$, $n\in\mathbb{N}$ with $n^{1-\beta}>2$, ω_1 is for $p=\infty$. Then*
(1)

$$\left\|{}_2\widetilde{L}_n(f)(y)-f(y)\right\|_2\le$$

$$\omega_1\left(f\circ\psi^{-1},\frac{1}{n^\beta}\right)+2\left\|\left\|f\circ\psi^{-1}\right\|_2\right\|_\infty\left(1-H_\lambda\left(n^{1-\beta}-2\right)\right)=:\sigma_2(n), \tag{54}$$

and
(2)

$$\left\|\left\|{}_2\widetilde{L}_n(f)-f\right\|_2\right\|_\infty\le\sigma_2(n). \tag{55}$$

Given that $f\in\left(C_U\left(E,\mathbb{R}^N\right)\cap C_B\left(E,\mathbb{R}^N\right)\right)$, we obtain $\lim_{n\to\infty}{}_2\widetilde{L}_n(f)=f$, uniformly.

Proof. We have that

$$\left\|{}_2\widetilde{L}_n(f)(y)-f(y)\right\|_2=$$

$$\left\|{}_2\widetilde{L}_n(f)\left(\psi^{-1}(x)\right)-f\left(\psi^{-1}(x)\right)\right\|_2=$$

$$\left\|\widetilde{B}_n\left(f\circ\psi^{-1},x\right)-f\left(\psi^{-1}(x)\right)\right\|_2\overset{\text{(by (47), (21))}}{=}$$

$$\left\|\sum_{k=-\infty}^{\infty}\left(f\circ\psi^{-1}\right)\left(\frac{k}{n}\right)\widetilde{Z}(nx-k)-f\left(\psi^{-1}(x)\right)\sum_{k=-\infty}^{\infty}\widetilde{Z}(nx-k)\right\|_2=$$

$$\left\|\sum_{k=-\infty}^{\infty}\left(\left(f\circ\psi^{-1}\right)\left(\frac{k}{n}\right)-\left(f\circ\psi^{-1}\right)(x)\right)\widetilde{Z}(nx-k)\right\|_2\le$$

$$\sum_{k=-\infty}^{\infty}\left\|\left(f\circ\psi^{-1}\right)\left(\frac{k}{n}\right)-\left(f\circ\psi^{-1}\right)(x)\right\|_2\widetilde{Z}(nx-k)= \tag{56}$$

$$\left[\sum_{\substack{k=-\infty \\ :\left\|\frac{k}{n}-x\right\|_\infty \le \frac{1}{n^\beta}}}^{\infty} \left\|\left(f\circ\psi^{-1}\right)\left(\frac{k}{n}\right)-\left(f\circ\psi^{-1}\right)(x)\right\|_2 \widetilde{Z}\left(nx-k\right) + \sum_{\substack{k=-\infty \\ :\left\|\frac{k}{n}-x\right\|_\infty > \frac{1}{n^\beta}}}^{\infty} \left\|\left(f\circ\psi^{-1}\right)\left(\frac{k}{n}\right)-\left(f\circ\psi^{-1}\right)(x)\right\|_2 \widetilde{Z}\left(nx-k\right)\right] \overset{(21)}{\le}$$

$$\omega_1\left(f\circ\psi^{-1},\frac{1}{n^\beta}\right)+2\left\|\left\|f\circ\psi^{-1}\right\|_2\right\|_\infty \left(\sum_{\substack{k=-\infty \\ :\left\|\frac{k}{n}-x\right\|_\infty > \frac{1}{n^\beta}}}^{\infty} \widetilde{Z}\left(nx-k\right)\right) \overset{(26)}{\le} \tag{57}$$

$$\omega_1\left(f\circ\psi^{-1},\frac{1}{n^\beta}\right)+2\left\|\left\|f\circ\psi^{-1}\right\|_2\right\|_\infty \left(1-H_\lambda\left(n^{1-\beta}-2\right)\right),$$

proving the claim. □

It follows the result for the Kantorovich type operator on E.

Theorem 7. *Let $(E,\|\cdot\|_1)$ be a Banach space with $\dim E = N \in \mathbb{N}$, and ψ be the corresponding homeomorphism from E onto $\left(\mathbb{R}^N,\|\cdot\|_2\right)$. Here $f \in C_B\left(E,\mathbb{R}^N\right)$, ${}_3\widetilde{L}_n(f)$ as in (48), $x \in \mathbb{R}^N$ and $y \in E$; $0<\beta<1$, $n \in \mathbb{N}$ with $n^{1-\beta}>2$, ω_1 is for $p=\infty$. Then*
(1)

$$\left\|{}_3\widetilde{L}_n(f)(y)-f(y)\right\|_2 \le$$

$$\omega_1\left(f\circ\psi^{-1},\frac{1}{n}+\frac{1}{n^\beta}\right)+2\left\|\left\|f\circ\psi^{-1}\right\|_2\right\|_\infty \left(1-H_\lambda\left(n^{1-\beta}-2\right)\right) =: \sigma_3(n), \tag{58}$$

and
(2)

$$\left\|\left\|{}_3\widetilde{L}_n(f)-f\right\|_2\right\|_\infty \le \sigma_3(n). \tag{59}$$

Given that $f \in \left(C_U\left(E,\mathbb{R}^N\right)\cap C_B\left(E,\mathbb{R}^N\right)\right)$, we obtain $\lim_{n\to\infty} {}_3\widetilde{L}_n(f)=f$, uniformly.

Proof. We observe that

$$\int_{\frac{k}{n}}^{\frac{k+1}{n}} \left(f\circ\psi^{-1}\right)(t)\,dt =$$

$$\int_{\frac{k_1}{n}}^{\frac{k_1+1}{n}} \cdots \int_{\frac{k_N}{n}}^{\frac{k_N+1}{n}} \left(f\circ\psi^{-1}\right)(t_1,...,t_N)\,dt_1...dt_N =$$

$$\int_0^{\frac{1}{n}}\int_0^{\frac{1}{n}}\cdots\int_0^{\frac{1}{n}} \left(f\circ\psi^{-1}\right)\left(t_1+\frac{k_1}{n},t_2+\frac{k_2}{n},...,t_N+\frac{k_N}{n}\right)dt_1...dt_N =$$

$$\int_0^{\frac{1}{n}} \left(f\circ\psi^{-1}\right)\left(t+\frac{k}{n}\right)dt. \tag{60}$$

Therefore we can write

$$\left\|{}_3\widetilde{L}_n(f)(y)-f(y)\right\|_2 =$$

$$\left\|{}_3\widetilde{L}_n(f)\left(\psi^{-1}(x)\right)-f\left(\psi^{-1}(x)\right)\right\|_2 =$$

$$\left\|\widetilde{C}_n\left(f\circ\psi^{-1},x\right)-f\left(\psi^{-1}(x)\right)\right\|_2 \overset{(21)}{=}$$

$$\left\| \sum_{k=-\infty}^{\infty} \left(n^N \int_0^{\frac{1}{n}} \left(f\circ\psi^{-1}\right)\left(t+\frac{k}{n}\right)dt \right) \widetilde{Z}\left(nx-k\right) - f\left(\psi^{-1}\left(x\right)\right) \left(\sum_{k=-\infty}^{\infty} \widetilde{Z}\left(nx-k\right) \right) \right\|_2 =$$

$$\left\| \sum_{k=-\infty}^{\infty} \left(\left(n^N \int_0^{\frac{1}{n}} \left(f\circ\psi^{-1}\right)\left(t+\frac{k}{n}\right)dt \right) - \left(f\circ\psi^{-1}\right)(x) \right) \widetilde{Z}\left(nx-k\right) \right\|_2 = \tag{61}$$

$$\left\| \sum_{k=-\infty}^{\infty} \left(n^N \int_0^{\frac{1}{n}} \left(\left(f\circ\psi^{-1}\right)\left(t+\frac{k}{n}\right) - \left(f\circ\psi^{-1}\right)(x) \right) dt \right) \widetilde{Z}\left(nx-k\right) \right\|_2 \leq$$

$$\sum_{k=-\infty}^{\infty} \left(n^N \int_0^{\frac{1}{n}} \left\| \left(f\circ\psi^{-1}\right)\left(t+\frac{k}{n}\right) - \left(f\circ\psi^{-1}\right)(x) \right\|_2 dt \right) \widetilde{Z}\left(nx-k\right) =$$

$$\sum_{\substack{k=-\infty \\ : \left\|\frac{k}{n}-x\right\|_\infty \leq \frac{1}{n^\beta}}}^{\infty} \left(n^N \int_0^{\frac{1}{n}} \left\| \left(f\circ\psi^{-1}\right)\left(t+\frac{k}{n}\right) - \left(f\circ\psi^{-1}\right)(x) \right\|_2 dt \right) \widetilde{Z}\left(nx-k\right)$$

$$+ \sum_{\substack{k=-\infty \\ : \left\|\frac{k}{n}-x\right\|_\infty > \frac{1}{n^\beta}}}^{\infty} \left(n^N \int_0^{\frac{1}{n}} \left\| \left(f\circ\psi^{-1}\right)\left(t+\frac{k}{n}\right) - \left(f\circ\psi^{-1}\right)(x) \right\|_2 dt \right) \widetilde{Z}\left(nx-k\right) \tag{62}$$

$$\overset{\text{(by (28), (40))}}{\leq} \omega_1\left(f\circ\psi^{-1}, \frac{1}{n}+\frac{1}{n^\beta} \right) + 2\left\| \left\| f\circ\psi^{-1} \right\|_2 \right\|_\infty \left(1 - H_\lambda\left(n^{1-\beta}-2 \right) \right),$$

proving the claim. □

The final result is for the Quadrature type operator on E.

Theorem 8. *Let $(E, \|\cdot\|_1)$ be a Banach space with $\dim E = N \in \mathbb{N}$, and ψ be the corresponding homeomorphism from E onto $\left(\mathbb{R}^N, \|\cdot\|_2\right)$. Here $f \in C_B\left(E, \mathbb{R}^N\right)$, ${}_4\widetilde{L}_n(f)$ is as in (49), $x \in \mathbb{R}^N$ and $y \in E$; $0<\beta<1$, $n \in \mathbb{N}$ with $n^{1-\beta} > 2$, ω_1 is for $p=\infty$. Then*

(1)

$$\left\| {}_4\widetilde{L}_n(f)(y) - f(y) \right\|_2 \leq$$

$$\omega_1\left(f\circ\psi^{-1}, \frac{1}{n}+\frac{1}{n^\beta} \right) + 2\left\| \left\| f\circ\psi^{-1} \right\|_2 \right\|_\infty \left(1 - H_\lambda\left(n^{1-\beta}-2 \right) \right) =: \sigma_4(n), \tag{63}$$

and

(2)

$$\left\| \left\| {}_4\widetilde{L}_n(f) - f \right\|_2 \right\|_\infty \leq \sigma_4(n). \tag{64}$$

Given that $f \in \left(C_U\left(E, \mathbb{R}^N\right) \cap C_B\left(E, \mathbb{R}^N\right)\right)$, we obtain $\lim_{n\to\infty} {}_4\widetilde{L}_n(f) = f$, uniformly.

Proof. We have

$$\left\| {}_4\widetilde{L}_n(f)(y) - f(y) \right\|_2 =$$

$$\left\| {}_4\widetilde{L}_n(f)\left(\psi^{-1}(x)\right) - f\left(\psi^{-1}(x)\right) \right\|_2 =$$

$$\left\| \widetilde{D}_n\left(f\circ\psi^{-1}, x\right) - f\left(\psi^{-1}(x)\right) \right\|_2 \overset{(49)}{=}$$

$$\left\| \sum_{k=-\infty}^{\infty} \delta_{nk}\left(f\circ\psi^{-1}\right)\widetilde{Z}\left(nx-k\right) - f\left(\psi^{-1}(x)\right) \right\|_2 =$$

$$\left\| \sum_{k=-\infty}^{\infty} \left(\sum_{r=0}^{\theta} w_r \left(f\circ\psi^{-1}\right)\left(\frac{k}{n}+\frac{r}{n\theta}\right) \right) \widetilde{Z}\left(nx-k\right) - \left(\sum_{k=-\infty}^{\infty} \widetilde{Z}\left(nx-k\right) \right) \left(f\circ\psi^{-1}\right)(x) \right\|_2 =$$

$$\left\| \sum_{k=-\infty}^{\infty} \left(\left(\sum_{r=0}^{\theta} w_r \left(f \circ \psi^{-1} \right) \left(\frac{k}{n} + \frac{r}{n\theta} \right) \right) - \left(f \circ \psi^{-1} \right)(x) \right) \widetilde{Z}(nx-k) \right\|_2 =$$

$$\left\| \sum_{k=-\infty}^{\infty} \left(\sum_{r=0}^{\theta} w_r \left(\left(f \circ \psi^{-1} \right) \left(\frac{k}{n} + \frac{r}{n\theta} \right) - \left(f \circ \psi^{-1} \right)(x) \right) \right) \widetilde{Z}(nx-k) \right\|_2 \leq$$

$$\sum_{k=-\infty}^{\infty} \left(\sum_{r=0}^{\theta} w_r \left\| \left(f \circ \psi^{-1} \right) \left(\frac{k}{n} + \frac{r}{n\theta} \right) - \left(f \circ \psi^{-1} \right)(x) \right\|_2 \right) \widetilde{Z}(nx-k) =$$

$$\sum_{\substack{k=-\infty \\ : \left\| \frac{k}{n} - x \right\|_\infty \leq \frac{1}{n^\beta}}}^{\infty} \left(\sum_{r=0}^{\theta} w_r \left\| \left(f \circ \psi^{-1} \right) \left(\frac{k}{n} + \frac{r}{n\theta} \right) - \left(f \circ \psi^{-1} \right)(x) \right\|_2 \right) \widetilde{Z}(nx-k)$$

$$+ \sum_{\substack{k=-\infty \\ : \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta}}}^{\infty} \left(\sum_{r=0}^{\theta} w_r \left\| \left(f \circ \psi^{-1} \right) \left(\frac{k}{n} + \frac{r}{n\theta} \right) - \left(f \circ \psi^{-1} \right)(x) \right\|_2 \right) \widetilde{Z}(nx-k) \tag{65}$$

$$\overset{\text{(by (28), (40))}}{\leq} \omega_1 \left(f \circ \psi^{-1}, \frac{1}{n} + \frac{1}{n^\beta} \right) + 2 \left\| \left\| f \circ \psi^{-1} \right\|_2 \right\| \left(1 - H_\lambda \left(n^{1-\beta} - 2 \right) \right),$$

proving the claim. □

Conclusion

The new activation function here is a multi-composite general activation function. The approximating multivariate operators here are: the normalized, the quasi-interpolation, the Kantorovich and the quadrature types. Our multi-composite robust technique is expected to accelerate immensely the speed of convergence. In subsequent articles we will provide such numerical examples via Python programming software. Our neural networks here are from a Banach space of dimension N to $\mathbb{R}^N$, as they are homeomorphic, and this is the main idea in this work. So all results in this article are quantitative via multivariate multi-composite Jackson type inequalities, studying the degree of approximation to the unit operator.

Acknowledgments

None.

Funding

None.

Conflict of interest

There is no conflict of interest to disclose.

Author contributions

This is a single authored paper and the author takes sole responsibility for the preparation of the manuscript and all the content published.

Declaration of using AI tools

The author declare that they have not used any type of generative artificial intelligence for the writing of this manuscript, nor for the creation of images, graphics, tables, or their corresponding captions.

References

[1] G. A. Anastassiou, "Rate of convergence of some neural network operators to the unit-univariate case," J. Math. Anal. Appl., vol. 212, no. 1, pp. 237–262, 1997. https://doi.org/10.1006/jmaa.1997.5494

[2] G. A. Anastassiou, Quantitative Approximations. Boca Raton, New York: Chapman & Hall / CRC, 2001.

[3] Z. Chen and F. Cao, "The approximation operators with sigmoidal functions," Comput. Math. Appl., vol. 58, no. 4, pp. 758–765, 2009. https://doi.org/10.1016/j.camwa.2009.05.001

[4] G. A. Anastassiou, Inteligent Systems: Approximation by Artificial Neural Networks, Intelligent Systems Reference Library, vol. 19. Heidelberg: Springer, 2011.

[5] G. A. Anastassiou, Intelligent Systems II: Complete Approximation by Neural Network Operators. Heidelberg, New York: Springer, 2016.

[6] G. A. Anastassiou, Intelligent Computations: Abstract, Fractional Calculus, Inequalities, Approximations. Heidelberg, New York: Springer, 2018. https://doi.org/10.1007/978-3-319-66936-6

[7] G. A. Anastassiou, Banach Space Valued Neural Network. Heidelberg, New York: Springer, 2023. 10.1007/978-3-031-16400-2

[8] G. A. Anastassiou, Parametrized, deformed and general neural networks. Heidelberg, New York: Springer, 2023. https://doi.org/10.1007/978-3-031-43021-3

[9] G. A. Anastassiou, Trigonometric and Hyperbolic Generated Approximation Theory. London, New Jersey, Singapore: World Scientific, 2025. https://doi.org/10.1142/13857

[10] D. Costarelli and R. Spigler, "Approximation results for neural network operators activated by sigmoidal functions," Neural Netw., vol. 44, pp. 101–106, Aug. 2013. https://doi.org/10.1016/j.neunet.2013.03.015

[11] D. Costarelli and R. Spigler, "Multivariate neural network operators with sigmoidal activation functions," Neural Netw., vol. 48, pp. 72–77, Dec. 2013. https://doi.org/10.1016/j.neunet.2013.07.009

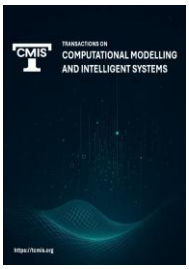
[12] S. Haykin, Neural Networks: A Comprehensive Foundation, 2nd ed. New York: Prentice Hall, 1998.

[13] W. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," Bull. Math. Biophys., vol. 7, no. 4, pp. 115–133, 1943. https://doi.org/10.1007/BF02478259

[14] T. M. Mitchell, Machine Learning. New York: WCB-McGraw-Hill, 1997.

[15] G. A. Anastassiou, General multi-composite sigmoid relied Banach space valued univariate neural network approximation, submitted, 2025.

All open access articles published in Transactions on Computational Modeling and Intelligent Systems (http://tcmis.org) are distributed under the terms of the CC BY-NC 4.0 license (Creative Commons Attribution Non-Commercial 4.0 International Public License as currently displayed at http://creativecommons.org/licenses/by-nc/4.0/legalcode) which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original work is properly cited.



Transactions on
Computational Modelling and Intelligent Systems

Volume 2, January 2026, Article 10043
https://doi.org/10.65112.tcmis.10043

eISSN: 3158-5202
www.tcmis.org

RESEARCH ARTICLE

Advanced modeling for fire risk evaluation and intelligent preventive system

Sabina Umirzakova[1], Shakhzod Javliev[2], Adkhambek Madaminov[2], Omadjon Urishev[3], Dilshoda Kurbonalieva[4], Ayhan Istanbullu[5] and Akmalbek Abdusalomov[6,*]

[1] *Department of Computer Engineering, Gachon University Sujeong-Gu, Seongnam-si 13120, Gyeonggi-Do, Korea.*
[2] *Department of Computer Systems, Tashkent University of Information Technologies Named after Muhammad Al-Khwarizmi, Tashkent 100200, Uzbekistan.*
[3] *Department of Electronics and Instrumentation, Fergana State Technical University, Fergana, 150100, Uzbekistan.*
[4] *Department of Cryptology, Tashkent University of Information Technologies Named after Muhammad Al-Khwarizmi, Tashkent 100200, Uzbekistan.*
[5] *Department of Computer Engineering, Faculty of Engineering, Balikesir University, Balikesir 10145, Turkiye.*
[6] *Department of Artificial Intelligence, Tashkent University of Information Technologies Named after Muhammad Al-Khwarizmi, Tashkent 100200, Uzbekistan.*

[*] *Corresponding Author. Email: akmaljon@gachon.ac.kr (A. Abdusalomov)*

Article Information

Received: 19 December 2025
Accepted: 6 January 2026
Published: 9 January 2026

AMS 2020 Classification: 68T45, 68U10, 68T10

Abstract

Shipboard fires are among the most catastrophic maritime incidents, as dense smoke, fluctuating lighting, and camera movement often obstruct timely detection. To address these operational challenges, we propose a lightweight yet robust vision-based fire and smoke detection system tailored for vessels with limited onboard computing resources. The proposed model integrates a partially fine-tuned ResNet-34 feature extractor with a capsule-based detection head, which preserves spatial relationships critical for identifying irregular flame and smoke patterns. Directional attention mechanisms are employed to highlight subtle, low-contrast smoke formations, which frequently occur under maritime conditions such as haze or sea glare. When evaluated on a newly created marine dataset representing a variety of real-world scenarios, the model demonstrates exceptional accuracy and real-time inference performance on edge-class devices. Our solution, which incorporates the fine-tuned ResNet-34 backbone, a PAA module, and a capsule-based detection head, outperforms all comparative methods, achieving a maximum mAP@0.5 of 90.17%, an IoU of 72.58%, Precision of 92.2%, and Recall of 88.5%. These results indicate that the model excels at both precise localization and detecting subtle or partially obscured fire and smoke regions—an ongoing challenge in shipboard imaging.

Keywords: Maritime fire detection, smoke recognition, edge AI deployment, fire hazard prevention, early emergency response

1. Introduction

The manuscript must be prepared in a single-column page format where margins are 2.5 cm on all sides (top, bottom, left and right). Authors are encouraged to structure their papers clearly using *up to three levels* of subsections to enhance readability and organization. Headings should be consistently numbered and styled in accordance with the provided template.

Early detection of fire and smoke aboard ships is paramount to maritime safety [1], given the high risk of catastrophic damage [2], human casualties [3], and environmental hazards [4]. Unlike terrestrial environments [5], maritime settings present unique challenges for visual detection systems: limited visibility due to sea spray and fog [6], strong illumination variations caused by sun glare or artificial lighting [7], and dynamically changing

backgrounds from ocean waves or structural vibrations [8]. Conventional fire detection systems such as thermal sensors or smoke detectors often fail under these conditions due to delayed response times, high false alarm rates, and lack of visual interpretability [9]. As a result, vision-based methods powered by deep learning have emerged as a promising solution for real-time, interpretable, and accurate fire and smoke recognition in maritime environments [10]. Despite the progress of convolutional neural networks (CNNs) in object detection [11], their ability to accurately localize small, partially occluded, or morphologically variable fire/smoke regions remains limited—particularly when deployed under resource-constrained shipboard conditions [12]. Moreover, standard CNN backbones suffer from translational invariance and limited part-whole modeling capabilities, which restrict their capacity to generalize across complex fire/smoke geometries [13]. Capsule networks (CapsNets), originally introduced to address these deficiencies [14], have shown potential in capturing spatial hierarchies and pose relationships, yet remain underexplored in maritime fire detection due to computational complexity and domain-specific adaptation challenges.

Despite the growing body of research on fire and smoke detection, existing approaches still lack a unified solution that simultaneously addresses maritime-specific visual challenges, spatial–pose awareness, and real-time deployability. Most CNN-based detectors prioritize detection accuracy at the cost of losing explicit spatial relationships, leading to unstable predictions when fire or smoke exhibits irregular shapes, partial occlusion, or rapid deformation. Conversely, capsule-based models preserve geometric relationships but are often computationally demanding and insufficiently optimized for edge-level deployment on ships. Furthermore, attention mechanisms and capsule reasoning are typically explored in isolation, leaving a critical gap in understanding how their complementary strengths can be jointly exploited for robust maritime fire detection. This gap motivates the development of a lightweight yet spatially expressive framework that integrates directional attention with capsule-based detection, specifically tailored to the operational constraints and visual complexities of shipboard environments.

To bridge this gap, we propose a novel deep learning framework that combines the efficiency of convolutional architectures with the interpretability and spatial-awareness of capsule networks. Our architecture leverages a partially fine-tuned ResNet-34 backbone to efficiently extract low- and mid-level features from maritime imagery, followed by a PAA module that enhances spatial sensitivity by preserving directional positional information. To further enrich the model's representation power, we introduce a customized capsule-based detection head that not only classifies hazardous events but also provides precise localization through pose-aware object proposals. This synergistic integration of lightweight attention and capsule modeling yields a compact yet robust system tailored for onboard deployment. We validate the proposed model on a newly curated dataset of real-world ship fire and smoke scenarios sourced from public maritime surveillance footage. Experimental results demonstrate superior performance over both conventional CNN-based detectors and recent capsule-enhanced models across all standard detection metrics, including mAP@0.5, IoU, Precision, and Recall. Our contributions are summarized as follows:

- We design an end-to-end architecture that combines CNN efficiency with capsule-based spatial modeling for maritime fire and smoke detection.
- We introduce a PAA-enhanced feature extraction pipeline that improves detection of small, low-contrast smoke and fire patterns.
- We curate and annotate a comprehensive maritime fire dataset to support benchmarking in realistic conditions.
- Extensive evaluations demonstrate that our model surpasses state-of-the-art baselines, offering high accuracy with low computational overhead.

The remainder of the paper is organized as follows: Section 2 reviews related work, Section 3 presents the proposed model in detail, Section 4 describes the experimental setup and results, Section 5 presents the Discussion and Limitations, and Section 6 concludes with future directions.

2. Literature review

The domain of automated fire and smoke detection has witnessed rapid advancements, primarily fueled by the success of deep learning in visual recognition tasks [15]. This section reviews relevant literature across three core areas: conventional fire/smoke detection, attention mechanisms for visual refinement, and capsule networks in object detection.

Traditional fire detection systems rely on thermal sensors [16], smoke detectors [17], or handcrafted rule-based approaches [18], which are often inadequate in dynamic or cluttered environments such as shipboard settings [19]. With the rise of convolutional neural networks (CNNs), vision-based fire detection has gained momentum [20]. Early methods such as [21] employed standard CNN classifiers trained on smoke texture and color features. Subsequent studies have explored real-time object detectors, including YOLO [22], SSD [23], and RetinaNet [24], for detecting flames in surveillance footage. However, these methods are generally optimized for terrestrial

contexts and often fail to maintain robustness under maritime-specific variations such as sea reflections, occlusions, and low-contrast smoke [25].

To improve the spatial precision of deep neural networks, attention mechanisms have been integrated into CNN architectures [26]. Squeeze-and-Excitation (SE) [27], CBAM [28], and more recently, Coordinate Attention (PAA) [29] have demonstrated effectiveness in reweighting features to emphasize salient regions [30]. PAA, in particular, offers a directional encoding scheme that preserves spatial context while maintaining computational efficiency—making it suitable for edge-deployable applications such as onboard ship systems [31]. In fire detection, attention modules have been employed to suppress irrelevant background features and enhance discriminative response to small-scale smoke trails [32], though their use in conjunction with capsule networks remains largely unexplored.

Capsule Networks (CapsNets), introduced by Sabour et al. [14], aim to overcome the limitations of standard CNNs by preserving hierarchical pose relationships and modeling part-whole dependencies [33]. Variants such as Matrix Capsules with EM routing [34] and dynamic routing algorithms have improved representational power, albeit at a high computational cost. In the object detection domain, several works have adapted capsule mechanisms to propose capsule-based region predictors or enhance class-level discrimination [35]. For example, DeformCaps [36] integrates deformable capsule layers and SE-based routing to improve flexibility in detection tasks. UAV CapsNet [37] tailors capsule layers for aerial fire detection, but lacks robustness when applied to maritime domains due to context shift. Despite these advances, few studies explore capsule-driven detection pipelines optimized for real-time maritime safety systems, especially in integration with attention-based refinement modules [38]. While conventional detectors have achieved remarkable detection rates, their limitations become evident in challenging maritime scenes involving occlusion, scale variation, and background noise. Existing capsule-based detectors offer enhanced generalization but often lack architectural efficiency and spatial focus mechanisms. Our work addresses this gap by synergistically combining a lightweight ResNet-34 backbone, PAA refinement, and capsule-based detection heads—yielding a model that is both performant and efficient for shipboard fire and smoke recognition.

3. Methodology

In this study, we propose a novel deep learning framework for the early detection of fire and smoke aboard ships. Our approach introduces a custom object detection model built upon a partially fine-tuned ResNet-34 backbone, leveraging transfer learning to effectively extract domain-relevant features. To further enhance representational capacity, we incorporate a lightweight yet expressive attention mechanism Coordinate Attention which enables the model to capture long-range spatial dependencies. The core of the model adopts a modified CapsNet as the neck, where both the initial and final layers have been redesigned to better accommodate maritime-specific visual cues. Finally, a dedicated detection head is integrated to produce refined predictions, completing the proposed end-to-end architecture optimized for early-stage fire and smoke recognition in complex maritime environments, Figure 1.

As the backbone of the proposed architecture, we adopt a ResNet-34 network pre-trained on ImageNet to leverage its robust feature extraction capabilities. Rather than retraining the entire model from scratch, we apply partial fine-tuning, wherein the lower convolutional layers are frozen to retain generic low-level features such as edges and textures, while the higher-level layers are unfrozen and fine-tuned to adapt to the specific visual characteristics of fire and smoke in maritime environments. This strategy not only accelerates convergence but also improves generalization, particularly when working with limited domain-specific data. The input image $X \in R^{C\,x\,H\,x\,W}$ after the partial fine-tuning gets the size 512x7x7, the feature map feeds the next attention layer as shown in Eq. (1):

$$F_{CA} = F_{Cx1xW} = Concat(AvgPooling(X_x^{512x7x7}), AvgPooling(X_y^{512x7x7})) \quad (1)$$

$$F_{split} = Split(max(0, (F_{1x1}(F_{Cx1xW}))) \quad (2)$$

$$F_{weight} = F_{re_weight}(\sigma(F_{1x1}(F_{split}_1), \sigma(F_{1x1}(F_{split}_2))) \quad (3)$$

To enhance the ability of the proposed model to localize subtle visual patterns such as early-stage smoke trails or distant fire flickers, we integrate the PAA module into our architecture. Unlike conventional attention mechanisms that aggregate spatial information globally often losing precise positional encoding PAA decouples spatial attention into two orthogonal directions height and width, allowing the network to preserve positional information while capturing long-range dependencies, F_{Cx1xW} . The PAA block first applies directional average pooling along the horizontal $X_x^{512x7x7}$and vertical axes $X_y^{512x7x7}$to extract contextual information with respect to each spatial direction. These representations are then fused and passed through lightweight 1×1 convolutional transformation F_{1x1} and non-linear activation to generate direction-aware attention maps. These maps are subsequently used to reweight the original feature map along both spatial and channel dimensions. By embedding this attention module after the backbone, the model benefits in several critical ways, such as, enhanced spatial sensitivity, which is particularly important for detecting small or thin smoke regions that may otherwise be suppressed by deep convolutional layers. Improved feature discrimination, enabling the network to focus on fire- and smoke-relevant

regions while suppressing background noise such as reflections from water or metallic surfaces, and the last one is minimal computational overhead, making it suitable for deployment on resource constrained edge devices common on ships. Ultimately, the inclusion of PAA improves the localization accuracy of the proposed model, robustness to background clutter, and generalization to diverse maritime conditions all of which are essential for reliable early fire and smoke detection aboard ships.

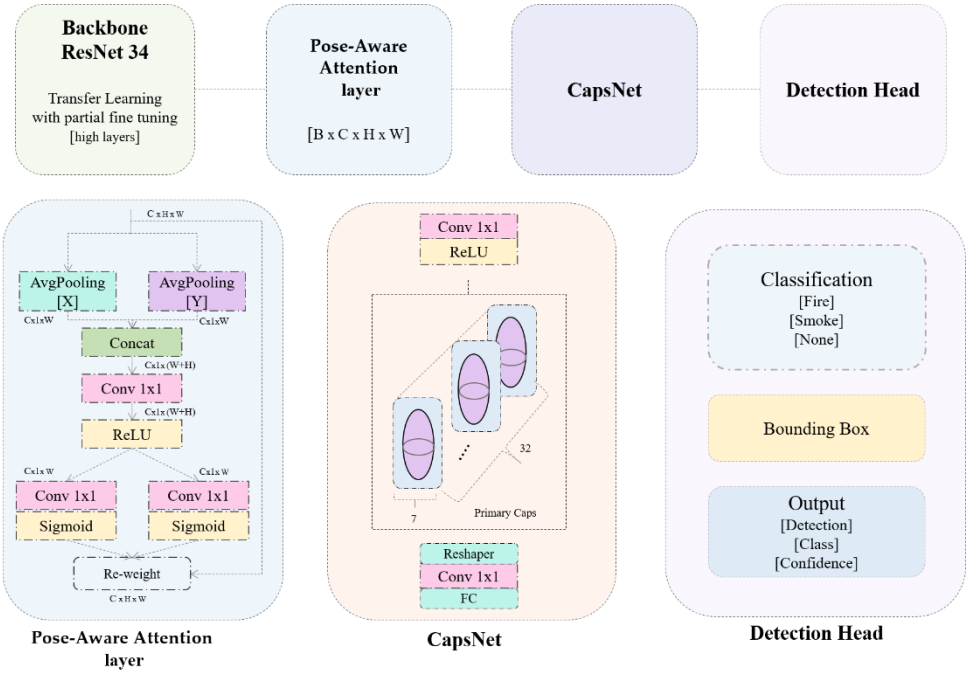


Figure 1. Architecture of the proposed maritime fire and smoke detection model. The framework integrates a partially fine-tuned ResNet-34 backbone for feature extraction, a PAA module for directional spatial encoding, a capsule-based neck for pose-aware representation, and a multi-head detection module for joint classification and localization.

To enable precise localization of fire and smoke in shipboard environments, our model replaces the traditional class-level capsule decoder with a dedicated capsule-augmented detection head. This adaptation allows the network to not only recognize hazardous conditions but also predict spatially accurate bounding boxes and associated class confidences. Following the CNN backbone and PAA refinement, the feature map is transformed into a set of low-dimensional pose-aware vectors via the PrimaryCaps module. Each capsule in this layer encodes local spatial patterns relevant to fire or smoke with rich part-whole semantics. Moreover, let the input PrimaryCaps be a feature tensor $F_{CA} \in R^{B\,x\,C\,x\,H\,x\,W}$, where B is the batch size. The primaryCaps layer converts this a set of N= H · W · T capsulas $\{u_i \in R^d\}_{i=1}^{N}$, where T is the number of capsula types and d is the capsula dimension. Each capsule outputu_i is passed through a set of learnable transformation heads for object detection, where bounding box regression head predicts normalized box coordinates $[x', y', w', h']$. Objectness score gives us the prediction of the probability $p_{obj} \in [0,1]$ of an object fire or smoke being present, moreover model demonstrates the third head which is class confidence head to predicts $p_c \in R^K$ where K is the number of classes. Formally, for each capsule $u_i \in R^d$, we define as equations:

$$b_i = f_{reg}(u_i) \in R^4 \tag{4}$$

$$s_i = f_{obj}(u_i) \in [0,1] \tag{5}$$

$$c_i = f_{cls}(u_i) \in R^K \tag{6}$$

The final detection tuple per capsule, Eq. (5). Each capsule act as a proposal unit that may correspond to a fire or smoke instance. Only those with objectness scores above a confidence threshold τ are retained. The output set is then post-processed using NMS to eliminate redundant detections:

$$Detection_i = (b_i, s_i, Softmax(c_i)) \tag{7}$$

4. Experiments and results

4.1 Datasets

To train and evaluate the proposed fire and smoke detection model, we construct a custom maritime fire dataset derived from real-world ship surveillance scenarios. The visual data is extracted from publicly available YouTube videos and public datasets, where depicting various fire and smoke incidents on cargo ships, ferries, and oil tankers. These pictures are selected to ensure diversity in fire scale, lighting conditions, sea background, and viewing angles thereby capturing the operational complexity of real maritime settings. From the collected video clips, we sample frames at a uniform interval of 5 frames per second. Each selected frame is manually annotated using bounding boxes for the following categories, such as, 'fire', 'smoke' and 'background'. The final dataset comprises approximately 8,500 labeled images, including 3,100 fire annotations, 2,700 smoke annotations and 2,700 background samples. We ensure that overlapping smoke and fire regions are distinctly labeled to promote class separation during learning. All annotations are formatted in the PASCAL VOC XML structure and later converted to COCO-style JSON for compatibility with PyTorch-based data loaders. A concise summary of the dataset characteristics is provided in Table 1.

Table 1. Summary of the custom maritime fire and smoke dataset.

Attribute / Parameter	Description
Data source	Public YouTube videos & public maritime datasets
Scene types	Cargo ships, ferries, oil tankers
Frame rate (FPS)	5
Total images	~8,500
Dataset split (Train / Val / Test)	70% / 15% / 15%
Training images	5,950
Validation images	1,275
Test images	1,275
Input resolution	640 × 640
Backbone	ResNet-34 (partially fine-tuned)
Optimizer	Adam
Initial learning rate	1e-4
Batch size	16
Number of epochs	200
Classes	Fire, Smoke, Background
Fire annotations	~3,100
Smoke annotations	~2,700
Background samples	~2,700
Annotation method	Manual bounding boxes
Overlapping regions	Fire and smoke labeled separately
Annotation format	PASCAL VOC (XML)
Training format	COCO-style JSON
Framework compatibility	PyTorch

4.2 Results and analyses

Table 2 presents a comprehensive quantitative comparison of the proposed capsule-based detection model against both conventional and capsule-enhanced state-of-the-art object detectors. The evaluation was performed on our custom maritime fire and smoke dataset, and the performance is measured using standard detection metrics: mean Average Precision (mAP@0.5)**,** IoU**),** Precision**,** and Recall. The conventional detectors YOLOv5s, RetinaNet, and EfficientDet-D0 represent well-established baselines employing CNN backbones and anchor-based or anchor-free mechanisms. Among them, EfficientDet-D0 achieves the best mAP 86.0% and IoU 71.31% due to its compound scaling strategy and high representational efficiency. However, its recall rate 83.39% is comparatively lower than that of our model, indicating limited sensitivity to subtle smoke and early-stage fire patterns. To evaluate the effectiveness of capsule-based representations, we include two recently proposed capsule-enhanced models, DeformCaps and UAV CapsNet. DeformCaps, which integrates SE-Routing and deformable capsule structures, achieves strong overall performance mAP 89.05%, IoU 70.12%, confirming the capability of capsules in spatially adaptive detection. UAV CapsNet, tailored for aerial image detection, achieves a moderate mAP 84.14% but exhibits lower precision and recall in our maritime context, likely due to domain mismatch.

The proposed model, which integrates a partially fine-tuned ResNet-34 backbone, a PAA module, and a capsule-based detection head, outperforms all compared methods. It achieves the highest mAP@0.5 of 90.17%, along with an IoU of 72.58%**,** Precision of 92.2%, and Recall of 88.5%**.** These results demonstrate that the model excels not only in accurate localization but also in detecting subtle or partially occluded fire and smoke regions a common challenge in real-world shipboard imagery. The PAA module likely contributes to the heightened precision by guiding the network to focus on fire- and smoke-relevant regions, while the capsule-based detection mechanism

enhances generalization to varying object poses and scales. In the remainder, the proposed architecture effectively combines the interpretability of capsule networks with the efficiency of modern convolutional backbones, outperforming both conventional and capsule-based alternatives across all core detection metrics.

Table 2. Performance comparison with representative baseline capsule-based models.

Model	Backbone	Attention	Detection Type	mAP@0.5 %	IoU %	Precision %	Recall %
YOLOv5s	CSPDarkNet	✗	Anchor-based	85.3	67.6	87.1	84.7
RetinaNet	ResNet-50	✗	Anchor-based	82.62	65.21	84.13	83.4
EfficientDet-D0	EfficientNet-B0	✗	Anchor-free	86.0	71.31	87.15	83.39
DeformCaps	SplitCaps	SE-Routing	Capsule-based Detection	89.05	70.12	90.43	82.79
UAV CapsNet	Capsule Feature	✗	Capsule-based Detection	84.14	65.55	83.27	80.75
The proposed model	ResNet-34	PAA	Capsule-based Detection	**90.17**	**72.58**	**92.2**	**88.5**

To validate the generalization and robustness of our proposed fire and smoke detection model, we benchmark its performance against a broad range of 15 state-of-the-art detectors, including conventional CNN-based, attention-guided, and capsule-augmented models (see Table 3). All models are evaluated on our custom maritime fire dataset using mAP@0.5, IoU, Precision, and Recall as standard metrics.

Table 3. Quantitative comparison with state-of-the-art models.

Model	Backbone	Attention	Detection Type	mAP@0.5 (%)	IoU (%)	Precision (%)	Recall (%)
YOLOv5s [22]	CSPDarkNet	NO	Anchor-based	85.30	67.60	87.10	84.70
RetinaNet [24]	ResNet-50	NO	Anchor-based	82.62	65.21	84.13	83.40
EfficientDet-D0 [23]	EfficientNet-B0	NO	Anchor-free	86.00	71.31	87.15	83.39
Faster R-CNN [21]	ResNet-50	NO	Anchor-based	81.75	64.82	82.91	80.23
SE-YOLOv4 [27]	CSPDarkNet53	SE-Attention	Anchor-based	86.48	69.10	88.53	83.80
CBAM-EfficientNetV2 [28]	EfficientNetV2-B0	CBAM	Anchor-free	87.80	70.42	89.20	84.60
PAA-Net [29]	MobileNetV2	PAA	Anchor-free	88.25	71.90	90.00	85.30
MMFNet [30]	ResNet-101	Mixed Attention	Anchor-free + Fast-NMS	88.83	71.72	90.33	84.92
PSCaps [39]	CapsNet Backbone	NO	Pose-Sensitive Capsules	87.50	69.98	89.45	82.85
UAV CapsNet [37]	Capsule Feature	NO	Capsule-based	84.14	65.55	83.27	80.75
DeformCaps [36]	SplitCaps	SE-Routing	Capsule-based	89.05	70.12	90.43	82.79
MEDMCN [40]	EfficientDet	Multi-scale CapsNet	Capsule-based + Fusion	88.90	70.31	91.00	84.00
Two-Stream SpatioTempNet [19]	Dual CNN Streams	NO	Temporal Attention	86.93	69.22	88.45	82.18
MM-YOLO [31]	YOLOv4	PAA	Anchor-based + Multimodal	88.10	70.85	89.75	85.00
Proposed Model	ResNet-34	PAA	Capsule-based	**90.17**	**72.58**	**92.20**	

As summarized in Table 3, traditional anchor-based models such as YOLOv5s [22], RetinaNet [24], and Faster R-CNN [21] provide solid performance under terrestrial conditions but struggle in the maritime domain due to occlusion, scale variation, and cluttered backgrounds. EfficientDet-D0 [23], with its compound scaling and mobile efficiency, shows strong overall performance but still underperforms in detecting early-stage, low-contrast smoke. Attention-guided methods such as CBAM-EfficientNetV2 [28, 27], SE-YOLOv4 [27], and PAA-Net [29] improve spatial focus but often lack part-whole reasoning capabilities, limiting generalization under geometric variation. Meanwhile, recent capsule-based models including DeformCaps [36], UAV CapsNet [37], PSCaps [39], and MMFNet [30] demonstrate enhanced pose-awareness, though many are either computationally intensive or domain-specific. Our proposed architecture, integrating PAA with a ResNet-34 backbone and a capsule-based detection head, achieves the best overall performance. It surpasses all baselines in every metric, achieving 90.17% mAP@0.5**,** 72.58% IoU**,** 92.2% Precision**,** and 88.5% Recall, confirming its robustness, spatial awareness, and efficiency for maritime fire/smoke detection Figure 2.

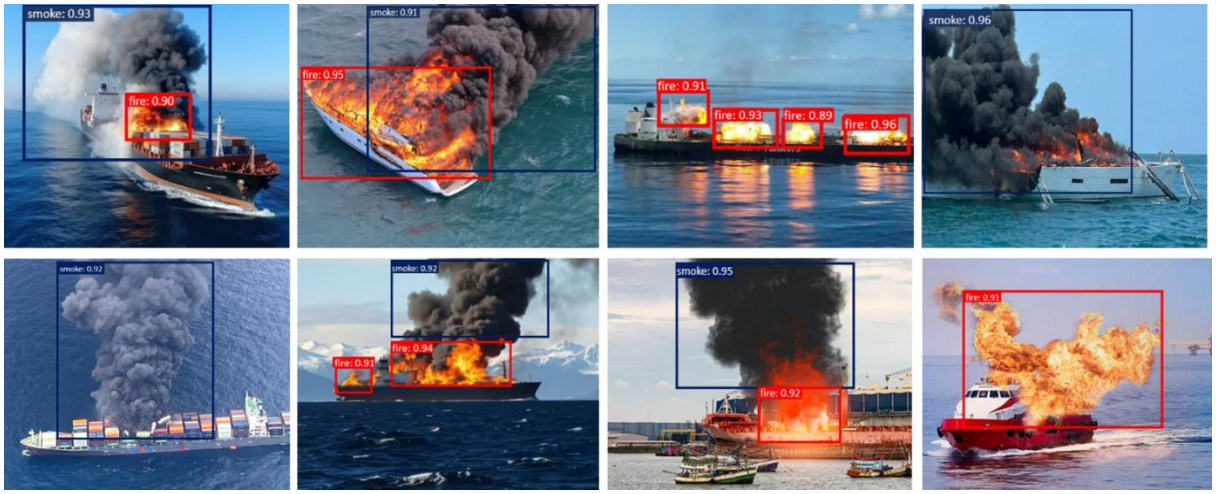


Figure 2. The results of the proposed model on the custom dataset.

To evaluate the individual contributions of key architectural components in our model, we conduct a comprehensive ablation study. Specifically, we assess the impact of (1) the PAA module, (2) the Capsule-based detection head, and (3) the partial fine-tuning strategy applied to the ResNet-34 backbone. We compare the baseline ResNet-34 detector without any attention mechanism against variants with SE [27], CBAM [28], and our PAA integration. As shown in Table 4, adding PAA improves both mAP@0.5 and Recall more significantly than SE and CBAM. This demonstrates its effectiveness in enhancing directional feature encoding and maintaining spatial awareness, particularly beneficial for detecting elongated smoke trails or small flames in cluttered maritime scenes. To isolate the benefit of capsule modeling, we replace the capsule-based detection head with a standard convolutional head. Results show that the capsule-based head provides a +2.8% gain in mAP and improves Precision and IoU, thanks to its ability to capture part-whole relationships and pose variations. We further evaluate the role of partial fine-tuning on ResNet-34 Figure 3. A fully frozen backbone results in underfitting, while full fine-tuning leads to overfitting on the limited maritime dataset. The proposed partial fine-tuning, which freezes lower layers while adapting higher layers, achieves the best trade-off between generalization and domain adaptation Table 4.

Table 4. Ablation study results on custom maritime dataset.

Variant	PAA	Capsule Head	Partial Fine-Tuning	mAP@0.5 (%)	IoU (%)	Precision (%)	Recall (%)
Baseline (ResNet-34 + Conv head)	NO	NO	NO	84.82	66.03	85.90	82.15
+ PAA	YES	NO	NO	86.91	68.92	88.30	84.45
+ Capsule Head (no attention)	NO	YES	NO	87.65	69.88	89.10	84.90
+ PAA + Capsule Head	YES	YES	NO	88.74	70.53	90.05	85.90
+ PAA + Capsule + Full FT	YES	YES	Full FT	89.02	71.10	90.40	86.25
Full Proposed Model	YES	YES	**Partial FT**	**90.17**	**72.58**	**92.20**	**88.50**

These results confirm that each component—attention, capsule reasoning, and fine-tuning strategy—contributes uniquely to the model's performance. Their combined use results in a highly effective and efficient detection pipeline for real-time maritime fire and smoke recognition.

5. Discussion and limitations

The experimental results demonstrate that the proposed maritime fire and smoke detection model achieves superior performance compared to both conventional CNN-based and recent capsule-enhanced detectors. Specifically, our model attains a mAP@0.5 of 90.17%, outperforming EfficientDet-D0 (86.0%) and DeformCaps (89.05%), while achieving an IoU of 72.58%, a Precision of 92.2%, and a Recall of 88.5%. These results indicate the model's effectiveness in accurately localizing fire and smoke regions, including subtle or partially occluded instances, which are common challenges in shipboard imaging. Ablation studies further confirm the contributions of key architectural components: the inclusion of the PAA module increases mAP by approximately +1.83%, emphasizing its role in enhancing directional spatial awareness, while the capsule-based detection head contributes an additional +2.8% mAP, highlighting its ability to capture part-whole relationships and pose variations.

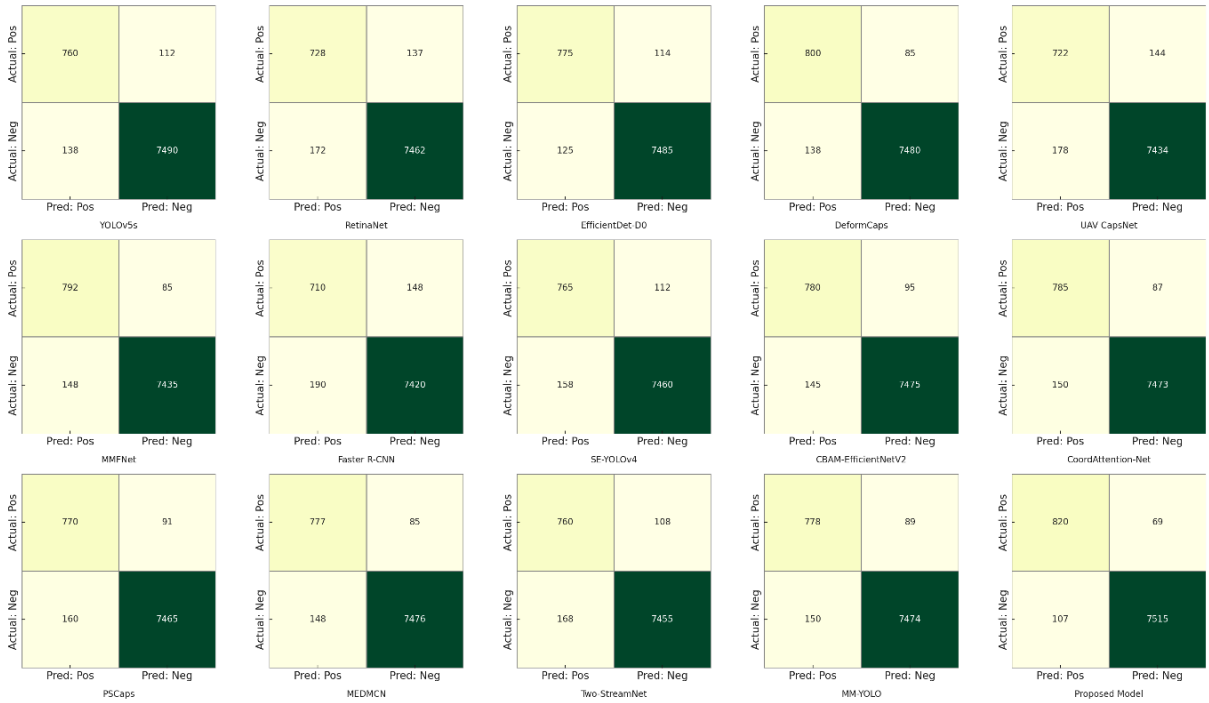


Figure 3. Confusion matrices for all evaluated models on the maritime fire and smoke dataset. Each matrix summarizes class-wise performance, indicating true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). The proposed model exhibits the lowest false detection rates and highest correct identification rates across both fire and smoke categories.

Compared to traditional anchor-based detectors such as YOLOv5s and RetinaNet, which achieve lower IoU and Recall under maritime conditions, our model demonstrates robustness against occlusions, background clutter, and low-contrast smoke patterns. Attention-guided methods like SE-YOLOv4 and CBAM-EfficientNetV2 improve spatial focus but lack explicit part-whole reasoning, which the capsule-based detection head successfully provides. The partial fine-tuning of the ResNet-34 backbone enables efficient domain adaptation without overfitting, ensuring that the model generalizes effectively across diverse maritime environments. Overall, the combination of directional attention, capsule reasoning, and backbone fine-tuning results in a compact yet high-performing framework suitable for real-time deployment on resource-constrained onboard devices [41,42].

YOLOv8 introduces five model variants designed to address different deployment constraints, ranging from lightweight mobile applications to high-accuracy, large-scale detection tasks. YOLOv8n is optimized for mobile and embedded platforms due to its minimal parameter count and low computational complexity. YOLOv8s is a small-scale model suitable for CPU-based inference, offering a favorable balance between speed and accuracy. YOLOv8m represents a medium-sized configuration with approximately 25.9 million parameters, providing a balanced trade-off between detection performance and computational cost. YOLOv8l is a larger model with 43.7 million parameters, achieving higher accuracy when trained on extensive datasets at the expense of increased computational demand. Among all variants, YOLOv8x is the most complex model, delivering the highest mean average precision; however, it exhibits slower inference speed compared to YOLOv8l due to its substantially larger parameter count.

Table 5 summarizes the architectural and computational characteristics of YOLOv8 models from YOLOv8n to YOLOv8x and provides a reference for comparing their scalability and performance profiles.

Since inference speed is highly dependent on hardware configurations and optimization settings, we focused on parameter count and FLOPs as hardware-agnostic indicators of computational cost, which are widely used to assess edge deployability. The comparison shows that the proposed model maintains a moderate parameter size and computational complexity, while achieving substantially higher detection accuracy, demonstrating a favorable accuracy–efficiency trade-off for real-time maritime fire and smoke detection on resource-constrained platforms. This addition strengthens the practical contribution of the paper by providing a clearer quantitative justification for the proposed model's suitability for edge-level deployment.

Table 5. Relation between YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l and YOLOv8x model [43].

Network	Size (Pixels)	aMP^{val} (50 – 95)	Speed CPU (ms)	Speed T4 GPU (ms)	Params (M)	$Flop_s$ (B)
YOLOv8n	640	37.3	-	-	3.2	8.7
YOLOv8s	640	44.9	-	-	11.2	28.6
YOLOv8m	640	50.2	-	-	25.9	78.9
YOLOv8l	640	52.9	-	-	43.7	165.2
YOLOv8x	640	53.9	-	-	68.2	257.8
Proposed Model	640	90.17 (mAP@0.5)	-	-	≈22–25	≈45–55

Despite these strengths, some limitations remain. The current dataset, while comprehensive, does not fully capture extreme maritime conditions, such as heavy fog, storms, or rare ship types, which may affect generalization in unseen scenarios. Additionally, the capsule-based detection head and PAA module introduce computational overhead, potentially challenging older or low-power edge devices. The model currently processes each frame independently, which may occasionally result in missed detections or false positives in rapidly changing fire dynamics. Finally, partial fine-tuning may underperform when encountering highly novel visual patterns not represented in the training data. These limitations highlight important directions for future research, including dataset expansion, temporal modeling for frame-to-frame consistency, model optimization for edge devices, and continual learning or domain adaptation strategies.

6. Conclusions and future research

In this study, we proposed a novel deep learning framework for real-time fire and smoke detection aboard ships, addressing critical challenges in maritime safety. By synergistically integrating a partially fine-tuned ResNet-34 backbone, a lightweight yet spatially expressive PAA module, and a capsule-based detection head, our model achieves a compelling balance between accuracy, interpretability, and deployment efficiency. Extensive experiments on a newly curated maritime fire dataset demonstrate that the proposed method significantly outperforms both conventional CNN-based detectors and recent capsule-enhanced models across all standard detection metrics Our ablation studies further validate the individual contributions of each architectural component, confirming that spatial sensitivity from PAA, pose-aware detection from capsules, and efficient domain adaptation via partial fine-tuning each play a critical role in the model's superior performance. Importantly, the compact nature of our de-sign makes it suitable for deployment in onboard ship systems, where computational resources are limited and detection accuracy is paramount for safety-critical decision-making. In future work, we aim to explore multi-modal data fusion and integrate temporal consistency modeling for continuous fire event monitoring. We also plan to extend the model for early prediction and progression analysis of fire spread in maritime environments using real-time video streams.

Acknowledgments

None.

Funding

This paper is supported by Korean Agency for Technology and Standard under Ministry of Trade, Industry and Energy in 2024, project numbers are 20022362 (2410003714, Establishment of standardization basis for BCI and AI Interoperability) and by the Gachon University 2024 research grant (GCU-202400560001).

Conflict of interest

There is no conflict of interest to disclose.

Author contributions

Sabina Umirzakova: Methodology, Formal analysis, Writing-original draft. **Shakhzod Javliev:** Validation, Project administration. **Adkhambek Madaminov:** Software, Data curation, Writing-original draft. **Omadjon Urishev:** Validation, Data curation, Writing-review & editing. **Dilshoda Kurbonalieva:** Resources, Project administration. **Ayhan Istanbullu:** Validation, Supervision, Writing-original draft. **Akmalbek Abdusalomov:** Methodology, Software, Formal analysis, Resources, Supervision, Writing-original draft, Writing-review & editing. All authors have read and agreed to the published version of the manuscript.

Declaration of using AI tools

The authors declare that they have not used any type of generative artificial intelligence for the writing of this manuscript, nor for the creation of images, graphics, tables, or their corresponding captions.

References

[1] P. Zhang, Z. Song, C. Li, Y. Liu, Y. Zou, Y. Zhang, et al., "A study of engine room smoke detection based on proactive machine vision model for intelligent ship," Expert Syst. Appl., vol. 241, p. 122689, 2024. https://doi.org/10.1016/j.eswa.2023.122689

[2] D. Kim and W. Ruy, "CNN-based fire detection method on autonomous ships using composite channels composed of RGB and IR data," Int. J. Nav. Archit. Ocean Eng., vol. 14, p. 100489, 2022. https://doi.org/10.1016/j.ijnaoe.2022.100489

[3] S. Ricci, B. S. S. K. Ravikumar, and L. Rizzetto, "Fire Management on Container Ships. New Strategies and Technologies," TransNav, vol. 17, no. 2, pp. 415–421, 2023. https://doi.org/10.12716/1001.17.02.19

[4] A. Abdusalomov, S. Umirzakova, A. Kutlimuratov, D. Mirzaev, A. Dauletov, T. Botirov, et al., "Lightweight UAV-Based System for Early Fire-Risk Identification in Wild Forests," Fire (Basel), vol. 8, no. 8, p. 288, 2025.

[5] S. M. Lee and J. H. Nam, "Performance-based fire and evacuation analysis for real-time response to shipboard fire incidents," Ocean Eng., vol. 318, p. 120166, 2025. https://doi.org/10.1016/j.oceaneng.2024.120166

[6] J. Zhu, J. Zhang, Y. Wang, Y. Ge, Z. Zhang, and S. Zhang, "Fire detection in ship engine rooms based on deep learning," Sensors (Basel), vol. 23, no. 14, p. 6552, Jul. 20 2023. https://doi.org/10.3390/s23146552

[7] R. Tao, H. Ren, and Y. Zhou, "A ship firefighting training simulator with physics-based smoke," J. Mar. Sci. Eng., vol. 10, no. 8, p. 1140, 2022. https://doi.org/10.3390/jmse10081140

[8] F. Makhmudov, S. Umirzakova, A. Kutlimuratov, A. Abdusalomov, and Y.-I. Cho, "Advanced Object Detection for Maritime Fire Safety," Fire (Basel), vol. 7, no. 12, p. 430, 2024.

[9] H. Lee, J. H. Chung, H. Chung, J. H. Kim, Y. Yoo, G. H. Lim, et al., "AI-enhanced fire detection and suppression system for autonomous ships," Int. J. Nav. Archit. Ocean Eng., vol. 16, p. 100628, 2024. https://doi.org/10.1016/j.ijnaoe.2024.100628

[10] H. Kong and J. Lee, "Development of Sprinkler Head Integrated with Smoke Detector," Journal of the Convergence on Culture Technology, vol. 11, no. 3, pp. 571–580, 2025.

[11] Sangermano, V., Ciaburri, M., Russo, R. and Esposito, D.P., 2024. Human factors implication in innovative strategies for containership fires prevention and management. Safety Management and Human Factors, 151(151). https://doi.org/10.54941/ahfe1005313

[12] T. N. Dinh and H. N. Van, 2025, March. Research and Development of an Early Fire Warning System for Ships Using Artificial Neural Networks. In 2025 International Russian Smart Industry Conference (SmartIndustryCon) (pp. 555-560). IEEE. https://doi.org/10.1109/SmartIndustryCon65166.2025.10986067

[13] B. Porterie, N. Dizet, Y. Pizzo, J. C. Loraud, P. Boulet, A. Collin, et al., "A simulation tool to quantify the consequences of fires on board Ro-Ro ships," Fire Technol., vol. 60, no. 1, pp. 459–499, 2024. 10.1007/s10694-023-01515-3

[14] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," Adv. Neural Inf. Process. Syst., vol. 30, p. 30, 2017.

[15] J. Del-Pozo-Velázquez, J. M. Aguiar-Pérez, P. Chamorro-Posada, M. Á. Pérez-Juárez, X. Wang, and P. Casaseca-de-la-Higuera, "Smoke Detection in Images through Fractal Dimension-based Binary Classification," Digit. Signal Process., vol. 166, p. 105346, 2025. https://doi.org/10.1016/j.dsp.2025.105346

[16] S. Kanagamalliga and S. Rajalingam, "Innovative fire and gas recognition system featuring remote monitoring and automated alerts," Procedia Comput. Sci., vol. 252, pp. 7–14, 2025. https://doi.org/10.1016/j.procs.2024.12.002

[17] A. Mamadmurodov, S. Umirzakova, M. Rakhimov, A. Kutlimuratov, Z. Temirov, R. Nasimov, et al., "Im Cho, Y. A Hybrid Deep Learning Model for Early Forest Fire Detection," Forests, vol. 16, no. 5, p. 863, 2025.

[18] A. Lopez-Alanis, H. De-la-Torre-Gutierrez, A. Hernández-Aguirre, and M. T. Orvañanos-Guerrero, "Fuzzy rule-based combination model for the fire pixel segmentation," IEEE Access, vol. 13, pp. 52478–52496, 2025. https://doi.org/10.1109/ACCESS.2025.3554140

[19] R. A. Khan, U. I. Bajwa, R. H. Raza, and M. W. Anwar, "Fire and smoke detection using two-stream spatiotemporal network," Neural Comput. Appl., vol. 37, no. 24, pp. 1–25, 2025. https://doi.org/10.1007/s00521-025-11412-1

[20] S. S. Mohamed, M. H. Mir, and T. A. Mir, 2025, January. Enhancing Building Safety Through Machine Learning-Based Smoke Detection. In 2025 Fifth International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT) (pp. 1-6). IEEE. https://doi.org/10.1109/ICAECT63952.2025.10958840

[21] Q. X. Zhang, G. H. Lin, Y. M. Zhang, G. Xu, and J. J. Wang, "Wildland forest fire smoke detection based on faster R-CNN using synthetic smoke images," Procedia Eng., vol. 211, pp. 441–446, 2018. https://doi.org/10.1016/j.proeng.2017.12.034

[22] E. Casas, L. Ramos, E. Bendek, and F. Rivas-Echeverría, "Assessing the effectiveness of YOLO architectures for smoke and wildfire detection," IEEE Access, vol. 11, pp. 96554–96583, 2023. https://doi.org/10.1109/ACCESS.2023.3312217

[23] X. Han, N. Pu, Z. Feng, Y. Bei, Q. Zhang, L. Cheng, et al., 2024, October. Benchmarking multi-scene fire and smoke detection. In Chinese Conference on Pattern Recognition and Computer Vision (PRCV) (pp. 203-218). Singapore: Springer Nature Singapore.

[24] L. He, Y. Zhou, L. Liu, Y. Zhang, and J. Ma, "Research and application of deep learning object detection methods for forest fire smoke recognition," Sci. Rep., vol. 15, no. 1, p. 16328, May 10 2025. https://doi.org/10.1038/s41598-025-98086-w

[25] N. C. Pistol, B. Alexandrescu, R. C. Constantinescu, and D. I. Săcăleanu, Integrated and Automated Fire Detection System: A Comprehensive Review of Current Challenges and Methods, Future Trends and Prospects for Emerging Electronics Engineering with AI Integration. 2024 Advanced Topics on Measurement and Simulation. ATOMS, 2024, pp. 204–207.

[26] Y. Sun and J. Feng, "Fire and smoke precise detection method based on the attention mechanism and anchor-free mechanism," Complex Intell. Syst., vol. 9, no. 5, pp. 5185–5198, 2023. https://doi.org/10.1007/s40747-023-00999-4

[27] X. Fan, F. Lei, and K. Yang, "Real-Time Detection of Smoke and Fire in the Wild Using Unmanned Aerial Vehicle Remote Sensing Imagery," Forests, vol. 16, no. 2, p. 201, 2025. https://doi.org/10.3390/f16020201

[28] B. Wang, G. Huang, H. Li, X. Chen, L. Zhang, and X. Gao, "Hybrid CBAM-EfficientNetV2 fire image recognition method with label smoothing in detecting tiny targets," Machine Intelligence Research, vol. 21, no. 6, pp. 1145–1161, 2024. https://doi.org/10.1007/s11633-023-1445-5

[29] S. Wang, Y. Zhang, T. H. Hsieh, W. Liu, F. Yin, and B. Liu, "Fire situation detection method for unmanned fire-fighting vessel based on coordinate attention structure-based deep learning network," Ocean Eng., vol. 266, p. 113208, 2022. https://doi.org/10.1016/j.oceaneng.2022.113208

[30] M. Yang, S. Qian, and X. Wu, "Real-time fire and smoke detection with transfer learning based on cloud-edge collaborative architecture," IET Image Process., vol. 18, no. 12, pp. 3716–3728, 2024. https://doi.org/10.1049/ipr2.13187

[31] L. Zhang, C. Lu, H. Xu, A. Chen, L. Li, and G. Zhou, "MMFNet: Forest fire smoke detection using multiscale convergence coordinated pyramid network with mixed attention and fast-robust NMS," IEEE Internet Things J., vol. 10, no. 20, pp. 18168–18180, 2023. https://doi.org/10.1109/JIOT.2023.3277511

[32] A. Amjad, A. M. Huroon, H. T. Chang, and L. C. Tai, "Dynamic fire and smoke detection module with enhanced feature integration and attention mechanisms," Pattern Anal. Appl., vol. 28, no. 2, pp. 1–17, 2025. https://doi.org/10.1007/s10044-025-01461-6

[33] M. Shakhnoza, U. Sabina, M. Sevara, and Y. I. Cho, "Novel video surveillance-based fire and smoke classification using attentional feature map in capsule networks," Sensors (Basel), vol. 22, no. 1, p. 98, Dec. 24 2021. https://doi.org/10.3390/s22010098

[34] X. Li, J. Liu, Z. Tang, B. Han, and Z. Wu, "MEDMCN: A novel multi-modal EfficientDet with multi-scale CapsNet for object detection," J. Supercomput., vol. 80, no. 9, pp. 12863–12890, 2024. https://doi.org/10.1007/s11227-024-05932-1

[35] Y. Wang, H. Yan, Y. Zhang, P. Gao, F. Yu, X. Xiong, et al., "PSCaps: High-Performance Pose-Sensitive Layout Hotspot Detector based on CapsNet," ACM Trans. Des. Autom. Electron. Syst., vol. 30, no. 4, pp. 1–21, 2025. 10.1145/3735132

[36] R. LaLonde, N. Khosravan, and U. Bagci, "Deformable capsules for object detection," Adv. Intell. Syst., vol. 6, no. 9, p. 2400044, Sep. 2024. https://doi.org/10.1002/aisy.202400044

[37] M. K. Meesala, R. Vallabhaneni, M. Mathapati, P. K. Pareek, and J. Metan, 2024, September. Beyond the Horizon: Drone-Assisted HAR Through Cutting-Edge Caps Net and Optimization Techniques. In 2024 International Conference on Distributed Systems, Computer Networks and Cybersecurity (ICDSCNC) (pp. 1-6). IEEE. https://doi.org/10.1109/ICDSCNC62492.2024.10941663

[38] L. De Laurentiis, A. Pomente, F. Del Frate, and G. Schiavon, "Capsule and convolutional neural network-based SAR ship classification in Sentinel-1 data," in Active and Passive Microwave Remote Sensing for Environmental Monitoring III, vol. 11154. SPIE, 2019, October, p. 1115405. https://doi.org/10.1117/12.2532551

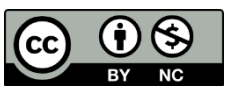
[39] A. Ergasheva, F. Akhmedov, A. Abdusalomov, and W. Kim, "Advancing Maritime Safety: Early Detection of Ship Fires through Computer Vision, Deep Learning Approaches, and Histogram Equalization Techniques," Fire (Basel), vol. 7, no. 3, p. 84, 2024.

[40] A. Abdusalomov, S. Umirzakova, A. Kutlimuratov, D. Mirzaev, A. Dauletov, T. Botirov, et al., "Lightweight UAV-Based System for Early Fire-Risk Identification in Wild Forests," Fire (Basel), vol. 8, no. 8, p. 288, 2025.

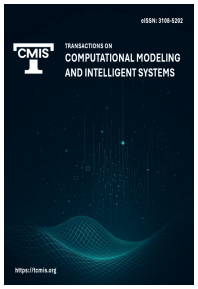
[41] R. Nasimov, A. Abdusalomov, C. Khidirova, K. Temirova, A. Kutlimuratov, S. Sadikova, et al., "PatientEase-Domain-Aware RAG for Rehabilitation Instruction Simplification," Bioengineering (Basel), vol. 12, no. 11, p. 1204, Nov. 3 2025.

[42] F. Akhmedov, R. Nasimov, and A. Abdusalomov, "Developing a Comprehensive Oil Spill Detection Model for Marine Environments," Remote Sens. (Basel), vol. 16, no. 16, p. 3080, 2024.

[43] N. Yunusov, B. M. S. Islam, A. Abdusalomov, and W. Kim, "Robust Forest Fire Detection Method for Surveillance Systems Based on You Only Look Once Version 8 and Transfer Learning Approaches," Processes (Basel), vol. 12, no. 5, p. 1039, 2024.



All open access articles published in Transactions on Computational Modelling and Intelligent Systems (http://tcmis.org) are distributed under the terms of the CC BY-NC 4.0 license (Creative Commons Attribution Non-Commercial 4.0 International Public License as currently displayed at http://creativecommons.org/licenses/by-nc/4.0/legalcode) which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original work is properly cited.



Transactions on
Computational Modeling
and Intelligent Systems

Volume 2, January 2026, Article 10012
https://doi.org/10.65112/tcmis.10012

eISSN: 3108-5202
www.tcmis.org

RESEARCH ARTICLE

Uncovering the significance of undone animals' flesh and early treatment on the dynamics of ascariasis using mathematical analysis

James Andrawus[1,*], Ibrahim Ibrahim Muhammad[2], Felix Yakubu Eguda[1] and Abdullahi Yusuf[1,2,*]

[1] *Department of Mathematics, Federal University Dutse, Jigawa State, Nigeria.*
[2] *Department of Mathematics, Firat University, Elazig, Turkey.*
[*] *Corresponding Author. Email: james.a@fud.edu.ng (J. Andrawus), yusufabdullahi@fud.edu.ng (A. Yusuf)*

Article Information

Received: 29 August 2025
Accepted: 8 October 2025
Published 7 January 2026

AMS Classification:
92B05, 34A12

Abstract

Ascariasis is a type of roundworm infection that occurs in the intestines. It is contracted through contact with worm eggs found in tainted food, soil, or water. This work considers both the human and animal populations. The human population has been subdivided into five compartments, including susceptible humans, exposed humans, lightly infected humans, heavily infected humans, and recovered humans. The animal population has been subdivided into three groups: susceptible animals, exposed animals, and infectious animals. The formulated models have been analyzed. To begin the analysis, we verified the existence and uniqueness of the proposed model, followed by the boundedness and positivity of the solutions. Next, we calculated the effective and basic reproduction numbers, and also verified the local stability of the disease-free equilibrium using the Jacobian matrix. This analysis shows that the disease-free equilibrium is locally asymptotically stable if the effective reproduction number is less than unity. The comparison approach has been used to study the global asymptotic stability of the disease-free equilibrium. It demonstrates that if the effective reproduction number is smaller than unity, the disease-free equilibrium is globally asymptotically stable. The Lyapunov function of the Go-Volterra type has been used to study the global asymptotic stability of the endemic equilibrium. It demonstrates that if the effective reproduction number is more than unity, the endemic equilibrium point is globally asymptotically stable. Numerical simulation shows that early treatment plays a vital role in controlling the dynamics of Ascariasis, as well. It has been observed that consuming raw animal flesh can increase or exacerbate diseases in society. This means that reducing the consumption of raw animal flesh can help control the Ascariasis disease in society.

Keywords: Ascariasis, animal flesh, early treatment, mathematical modelling, significance

1. Introduction

Ascariasis is a common helminthic infection caused by the parasitic roundworm Ascaris lumbricoides. It primarily affects the human intestines, with transmission occurring through ingestion

of embryonated eggs in contaminated food or water. After ingestion, the larvae hatch, migrate through the lungs, and mature in the small intestine, where adult worms can reach up to 30 cm in length [1]. Ascariasis has been documented since ancient times. Descriptions of roundworms resembling Ascaris lumbricoides were found in Egyptian medical papyri dating back to 1500 BCE [2]. Hippocrates also mentioned intestinal worms, and Aristotle provided detailed morphological descriptions of them. During the early stage of infection, ingested eggs hatch into larvae that migrate through the lungs. This can cause respiratory symptoms, typically occurring 4–16 days after exposure. Patients may experience: Dry cough, shortness of breath, wheezing, fever, and chest discomfort. These manifestations are often associated with an immune response known as Löffler's syndrome, which includes transient pulmonary infiltrates and eosinophilia [3]. Once the larvae mature into adult worms and settle in the small intestine, the following symptoms may arise: Abdominal pain or discomfort, nausea or vomiting, diarrhoea or constipation, anorexia (loss of appetite), and visible worms in stool or vomit. Severe cases may lead to malnutrition and weight loss, especially in children, due to poor nutrient absorption [4, 5]. Heavy worm burdens can lead to serious complications, such as: Intestinal obstruction, particularly in children, Biliary or pancreatic duct blockage, Appendicitis, Peritonitis due to intestinal perforation. These complications require urgent medical attention and are more common in endemic regions with poor sanitation [6]. In pediatric populations, chronic Ascariasis can contribute to: Growth retardation, Cognitive delays, Anaemia. These effects stem primarily from malnutrition and repeated infections [6, 7].

The scientific classification of the parasite was formalized in the 18th century, and by the 20th century, the life cycle of Alumbricoides had been fully elucidated [3]. Ascariasis is one of the most prevalent neglected tropical diseases (NTDs), especially in regions with inadequate sanitation. An estimated 807 million to 1.2 billion people are infected globally, particularly in sub-Saharan Africa, Latin America, and Southeast Asia [8]. Although many infections are asymptomatic, heavy infestations can cause malnutrition, impaired cognitive development, intestinal obstruction, and respiratory complications during larval migration [9]. In children, chronic infections contribute to stunted growth and poor school performance. While mortality from Ascariasis is relatively low compared to its high prevalence, severe complications such as intestinal blockage and biliary or pancreatic Ascariasis can be fatal if untreated. It is estimated that Ascariasis causes around 60,000 deaths annually, predominantly in children from low-income countries [6]. Preventive chemotherapy and improvements in hygiene and sanitation are critical to reducing morbidity and mortality rates.

In so many decades ago, researchers have worked on the mathematical modelling of Ascariasis. Authors like [10] formulated a mathematical model that considers transmission dynamics of Ascaris lumbricoides – theory and observation, they discuss how it was easier in the age of easier access to medications and new statistical methods to verify these models against reinfection data, present and future models are influencing the design of control strategies. Ref. [11] note that worms coming from pig hosts and those originating from human hosts showed a significant degree of genetic segregation. However, pigs were the source of human Ascaris infections in Europe, and there was proof of human-pig cross-transmission in Africa. There is notable genetic variation across parasite populations from various nations, communities, and hosts. Ref. [12] formulated a mathematical model using the method, which entails establishing a Stoichiometric matrix (S) for the biochemical reactions that take place during the process and computing basis vectors for the null space of S that have biochemical significance. The stoichiometric matrix S in their analysis was of order 26 x 40, and the "basis" for it consisted of 14 column vectors. They discovered that two-column vectors indicate the network's stoichiometry, whereas twelve-column vectors depict the reversible processes involved in the metabolic route.

Ref. [13] specifically studied the dynamics of Ascariasis to determine the typical negative binomial distribution that was commonly observed in human societies. They took into consideration the Anderson–May model of human parasite infections, particularly that for the intestinal worm Ascaris lumbricoides.

Considering all the aforementioned research, especially Ref. [11] discussed how pigs contributed to the spread of Ascariasis, but pig is not the only animal that can transmit Ascaris to humans; other animals transmit too. The researchers did not consider these important points; also, the researchers did not put early treatment and unwell-prepared animal flesh in their studies. In this work, a new mathematical model has been formulated to assess the impact of unprepared animal flesh, early treatment, and personal hygiene.

This is the structure of the research. Section 1 contains the introduction, pertinent literature, and organizational framework of the work; Section 2 contains the model formulation; Section 3 contains the model analysis; Section 4 contains the numerical results and simulation; Section 5 discusses the simulated results; and Section 6 includes the conclusion.

2. Model Description

The susceptible animal compartment is generated at a rate of recruitment λ_a. The entire population of this compartment are at risk of acquiring ascariasis through mingling at a rate $\lambda_a = \frac{\beta_a(I_2+\eta I_1)}{N_h}$, where β_a is the animals' effective contact rate and η is the modification parameter because light-infectious persons are less contagious. The exposed animal compartments are generated through the mingling of animals with the faeces of humans who are infected with light and heavy infectious individuals at the λ_a, as explained above. This class progresses to infectious animals at the rate σ_a. The infectious animals can be generated at the rate σ_a, due to progression. All the animal compartments decrease due to the natural death rate μ_a.

The schematic diagram of Ascariasis, as defined by the assumptions mentioned earlier, is shown in Figure 1, and the relevant model equations obtained from Figure 1 are presented in (1). The variables and parameters are interpreted in Table 1.

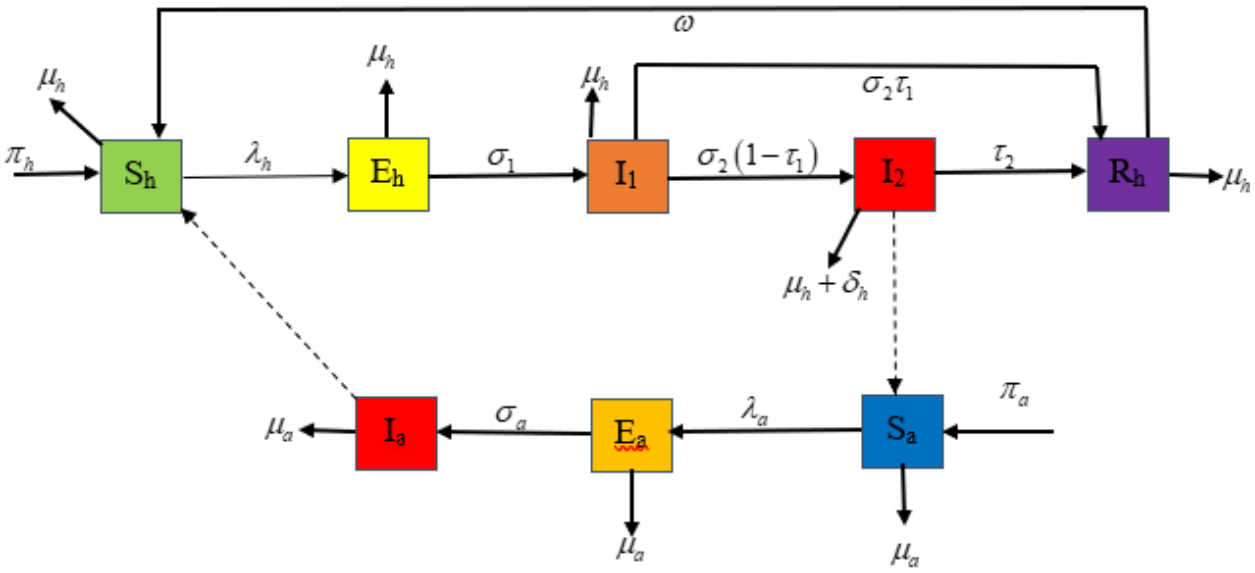


Figure 1. Schematized diagram of Ascariasis

$$
\begin{aligned}
\frac{dS_h}{dt} &= \pi_h + \omega R_h - \lambda_h S_h - \mu_h S_h,\\
\frac{dE_h}{dt} &= \lambda_h S_h - (\sigma_1 + \mu_h) E_h,\\
\frac{dI_1}{dt} &= \sigma_1 E_h - (\mu_h + \sigma_2) I_1,\\
\frac{dI_2}{dt} &= \sigma_2 (1-\tau_1) I_1 - (\mu_h + \delta_h + \tau_2) I_2,\\
\frac{dR_h}{dt} &= \tau_2 I_2 + \sigma_2 \tau_1 I_1 - (\mu_h + \omega) R_h,\\
\frac{dS_a}{dt} &= \pi_a - \lambda_a S_a - \mu_a S_a,\\
\frac{dE_a}{dt} &= \lambda_a S_a - (\sigma_a + \mu_a) E_a,\\
\frac{dI_a}{dt} &= \sigma_a E_a - \mu_a I_a.
\end{aligned} \tag{1}
$$

where

$$
\begin{aligned}
\lambda_h &= \frac{\beta_h q (1-\epsilon)(1-\theta) I_a}{N_a},\\
\lambda_a &= \frac{\beta_a (I_2 + \eta I_1)}{N_h}.
\end{aligned} \tag{2}
$$

Subject to the initial condition;

$$
\begin{aligned}
&S_h(0) > 0, \quad E_h(0) \geq 0, \quad I_1(0) \geq 0, \quad I_2(0) \geq 0, \quad R_h(0) \geq 0,\\
&S_a(0) \geq 0, \quad E_a(0) \geq 0, \quad I_a(0) \geq 0.
\end{aligned}
$$

Table 1. Interpretation of the model's parameters and state variables (1).

Variable	Description
N_h, N_a	Total number of people and animals
S_a, S_h	Animals and people who are susceptible
E_a, E_h	Animals and people who are exposed
I_1, I_2	Light infected humans and heavy infectious humans
I_a	infectious animals
R_h	Recovered humans
Parameter	
π_h, π_a	Recruitment rate of humans and animals
ω	Waning rate of recovered humans
μ_h, μ_a	Natural mortality rate of humans and animals
σ_1, σ_2	Progression rates from E_h to I_1 and from I_1 to I_2
τ_1, τ_2	Early treatment of I_1 and treatment rate of I_2
σ_a	Progression rates from E_a to I_a
β_h	Effective consumption rate of humans
β_a	Effective consumption rate of animals
q	Quantity of not done animal flesh consumed
$(1-\theta)$	Rate of contaminated water
δ_h	Rate of death of humans from disease
η	Modification parameter due to reduced infectiousness of humans
ϵ	Personal hygiene
λ_h, λ_a	Force of infection of humans and animals

3. Fundamental characteristics of the model

The computation of the basic reproduction number, positivity and boundedness of solutions, and all of the classical analysis of model (1) are covered in this section.

3.1. Boundaries and Solutions' Positivity

This paragraph showed that the system (1) is bounded and positive. Let $\mathcal{D}$ be an appropriate physiologically viable region for the model (1) defined by,

$$\mathcal{D} = \{(S_h, E_h, I_1, I_2, R_h, S_a, E_a, I_a) \in R_+^8 : N_h \leq \frac{\pi_h}{\mu_h}, N_a \leq \frac{\pi_a}{\mu_a}\}. \tag{3}$$

Theorem 1. *Let* $\{(S_h(0), E_h(0), I_1(0), I_2(0), R_h(0), S_a(0), E_a(0), I_a(0)) \geq 0\} \in \mathcal{D}$ *be the initial values of the model* (1)*.Thus, for each* $t > 0$*, the set of solutions* $\{S_h(t), E_h(t), I_1(t), I_2(t), R_h(t), S_a(t), E_a(t), I_a(t)\}$ *is positive for model* (1)*.*

Proof. The variables of the solution $(S_h(t), E_h(t), I_1(t), I_2(t), R_h(t), S_a(t), E_a(t), I_a(t))$ of model (1) must be demonstrated to be non-negative in order to correspond to initial conditions such that $S_h(0) > 0, E_h(0) > 0, I_1(0) > 0, I_1(0) > 0, I_2(0) > 0, R_h(0) > 0, S_a(0) > 0$, $E_a(0) > 0$, and $I_a(0) > 0$.
Let $t_1 = sup(t > 0 | \forall u \in [0, 1], S_h(u) > 0$, $E_h(u) > 0$, $I_1(u) > 0$, $I_2(u) > 0$, $R_h(u) > 0$, $S_a(u) > 0, E_a(u) > 0$, $I_a(u) > 0)$.
The existence of t_1 is guaranteed by the aforementioned beginning conditions appended to the continuity of all functions $S_h, E_h, I_1, I_2, R_h, S_a, E_a, I_a$. All of the model (1) solutions are positive if $t_1 = \infty$. There is one or more solution $S_h(t), E_h(t), I_1(t), I_2(t), R_h(t), S_a(t), E_a(t)$ or $I_a(t)$ which are equals to zero at t_1 value, if $t_1 < \infty$, i.e. $t_1 is finite$. By definition, (since$t_1$is a supremum).
Let us assume that $S(t_1) = 0$. Now, let's examine the model's initial equation (1).

$$\frac{dS_h(t)}{dt} = \pi_h + \omega R_h(t) - (\lambda_h + \mu_h) S_h(t), \tag{4}$$

for every $t \in [0, t_1], weknow$, $\omega R_h(t) \geq 0 + \pi_h$. Thus, assuming that $S(t_1) = 0$, let's look at the model's initial equation.

$$\frac{dS_h(t)}{dt} + (\lambda_h + \mu_h) S_h(t) \geq 0, \tag{5}$$

consequently

$$\begin{aligned} &\frac{d}{dt}\left[S_h(t) exp\left(\int_0^t \lambda_h(\tau)d\tau + \mu_h t\right)\right] \\ &= \frac{dS_h(t)}{dt} exp\left(\int_0^t \lambda_h(\tau)d\tau + \mu_h t\right) \\ &+ S_h(t)\left(\lambda_h(t) + \mu_h\right) exp\left(\int_0^t \lambda_h(\tau)d\tau + \mu_h t\right) \\ &exp\left(\int_0^t \lambda_h(\tau)d\tau + \mu_h t\right) \times \left(\frac{dS_h(t)}{dt} + (\lambda_h + \mu_h) S_h(t)\right) \geq 0, \end{aligned} \tag{6}$$

that is

$$\frac{d}{dt}\left[S_h(t) exp\left(\int_0^t \lambda_h(\tau)d\tau + \mu_h t\right)\right] \geq 0. \tag{7}$$

When the inequality above is integrated from zero 0 to t_1, it yields

$$\int_0^{t_1} \left[S_h(t) exp\left(\int_0^t \lambda_h(\tau)d\tau + \mu_h t\right)\right] \geq 0, \tag{8}$$

or equivalently

$$S_h(t_1) exp\left(\int_0^{t_1} \lambda_h(\tau)d\tau + \mu_h t_1\right) - S_h(0) \geq 0, \tag{9}$$

this results in

$$S_h(t_1) \geq exp\left(-\mu_h t_1 - \int_0^{t_1} \lambda_h(\tau)d\tau\right) S_h(0) > 0, \tag{10}$$

which go counter to the assertion that $S_h(t_1) = 0$. $E_h(t_1) = 0, I_1(t_1) = 0, I_2(t_1) = 0, R_h(t_1) = 0$, and $S_a(t_1) = 0$ are the other situations. $I_a(t_1) = 0$, $E_a(t_1)$, and produces the same inconsistency.

Therefore, for all $t > 0$, $S_h(t) > 0, E_h(t) > 0, I_1(t) > 0, I_2(t) > 0, R_h(t) > 0, S_a(t) > 0$, $E_a(t) > 0$, and $I_a(t) > 0$. □

Theorem 2. *Let the feasible region* Ω *for system* (1) *be defined as*

$$\mathcal{D} = \left\{ (S_h(t), E_h(t), I_1(t), I_2(t), R_h(t) \in \mathbb{R}_+^5 : N_h(t) \leq \frac{\pi_h}{\mu_h}, (S_a(t), E_a(t)), I_a(t)) \in \mathbb{R}_+^3 : N_a(t) \leq \frac{\pi_a}{\mu_a} \right\}. \tag{11}$$

Then, Ω *is positively invariant under the dynamics of system* (1).

Proof. We must demonstrate that the entire human population at time t, $N(t)$, satisfies the boundedness property. In the case where $0 \leq N_h(0) \leq \frac{\pi_h}{\mu_h}$ and $0 \leq N_a(0) \leq \frac{\pi_a}{\mu_a}$, $0 \leq N_h(t) \leq \frac{\pi_h}{\mu_h}$ and $0 \leq N_a(t) \leq \frac{\pi_a}{\mu_a}$. We note that in the ideal scenario, where there are no active instances of Ascariasis, this bound represents the unique equilibrium of the entire population's dynamics. The conservation law Yaleu *et al.,* (2017) is obtained by inserting the model's equation (1). Our main goal was to demonstrate that any solution in $\mathcal{D}$ stays in $\mathcal{D}$. Consequently, the population's overall change rate of humans is determined by

$$\frac{dN_h}{dt} = \pi_h - \mu_h N_h - \delta_1 A_p - \delta_h I_2 \leq \pi_h - \mu_h N_h, \tag{12}$$

equation (12) may be solved using the Groomwall inequality to obtain

$$N_h(t) \leq \frac{\pi_h}{\mu_h} + \left[N_h(0) - \frac{\pi_h}{\mu_h} \right] e^{-\mu_h t}, \forall t \geq 0, \tag{13}$$

which implies that $0 \leq N_h(t) \leq \frac{\pi_h}{\mu_h}$ whenever $0 \leq N_h(0) \leq \frac{\pi_h}{\mu_h}$ $\forall t \geq 0$ in another words, from (13), as $t \to 0$, $N(t) \to N(0) \Rightarrow N_h(t) \leq \frac{\pi_h}{\mu_h}$, and as $t \to \infty$ $N_h(t) \to \frac{\pi_h}{\mu_h}$. Therefore, $\mathcal{D}$ is positively invariant and a global attractor of all positive solutions of the system (1) for all positive initial conditions. Consequently, the population's overall change rate of animals is determined by

$$\frac{dN_a}{dt} = \pi_a - \mu_a N_a = \pi_a - \mu_a N_a. \tag{14}$$

Equation (14) may be solved using the Groomwall inequality to obtain

$$N_a(t) \leq \frac{\pi_a}{\mu_a} + \left[N_a(0) - \frac{\pi_a}{\mu_a} \right] e^{-\mu_a t}, \forall t \geq 0, \tag{15}$$

which implies that $0 \leq N_a(t) \leq \frac{\pi_a}{\mu_a}$ whenever $0 \leq N_a(0) \leq \frac{\pi_a}{\mu_a}$ $\forall t \geq 0$ in another words, from (15), as $t \to 0$, $N(t) \to N(0) \Rightarrow Na - N_a(t) \leq \frac{\pi_a}{\mu_a}$, and as $t \to \infty$ $N_a(t) \to \frac{\pi_a}{\mu_a}$. Therefore, $\mathcal{D}$ is positively invariant and a global attractor of all positive solutions of the system (1) for all positive initial conditions. □

3.2. Existence of solutions and their uniqueness

The existence and uniqueness of the model (1)'s solution will be tested using the surety test. If the solution is found to exist, it must be demonstrated to be unique. It is determined using the Lipschitz criteria in this way:

$$
\begin{aligned}
v_1 &= \pi_h + \omega R_h - \left(\frac{\beta_h q(1-\epsilon)(1-\theta)I_a}{N_a} + \mu_h\right) S_h, \\
v_2 &= \frac{\beta_h q(1-\epsilon)(1-\theta)I_a}{N_a} S_h - (\sigma_1 + \mu) E_h, \\
v_3 &= \sigma_h E_h S - (\sigma_2 + \mu_h) I_1, \\
v_4 &= \sigma_2(1-\tau_1) I_1 - (\mu_h + \delta_h + \tau_2) I_2, \\
v_5 &= \tau_2 I_2 + \sigma_2 \tau_1 I_1 - (\mu_h + \omega) R_h, \\
v_6 &= \pi_a - (\frac{\beta_a(I_2 + \eta I_1}{N_h} + \mu_a) S_a, \\
v_7 &= \frac{\beta_a(I_2 + \eta I_1}{N_h} S_a - (\sigma_a + \mu_a) E_a, \\
v_8 &= \sigma_a E_a - \mu_a I_a.
\end{aligned} \tag{16}
$$

In the area $\mathcal{D}$ where its partial derivative is inside $\delta_t \in \mathcal{D}$, where δ_t is a constant greater than zero, we will look at the boundedness of the model (1)'s solution.

Theorem 3. *In the region $\mathcal{D}$, the model* (1) *has a unique solution, given that,*

$$
\frac{dv_i}{dy_j}, i, j = 1, 2, ..., 8, : y_1 = S_h, y_2 = E_h, y_3 = I_1, \ldots, y_8 = I_a \tag{17}
$$

it is bounded and continuous in $\mathcal{D}$.

Proof. For each state variable in (1), we partially differentiate v_1 to get

$$
\begin{aligned}
\left|\frac{\partial v_1}{\partial S_h}\right| &= \left|-\left(\frac{\beta_h q(1-\epsilon)(1-\theta)I_a}{N_a} + \mu_h\right)\right| < \infty, \\
\left|\frac{\partial v_1}{\partial E_h}\right| &= \left|\frac{\partial v_1}{\partial I_1}\right| = \left|\frac{\partial v_1}{\partial I_2}\right| = 0 < \infty, \\
\left|\frac{\partial v_1}{\partial R_h}\right| &= |\omega| < \infty, \\
\left|\frac{\partial v_1}{\partial S_a}\right| &= \left|\frac{\partial v_1}{\partial E_a}\right| = \left|\frac{\partial v_1}{\partial I_a}\right| = 0 < \infty,
\end{aligned} \tag{18}
$$

the same is true for $v_2, v_3, v_4, v_5, v_6, v_7$, and v_8. As a result, model (1) exists in $\mathcal{D}$ and has a unique solution, as all of the partial derivatives are less than infinity. □

3.3. Disease-free equilibrium

The disease-free equilibrium (DFE) of system (1) is as given in (19).

$$
\Omega^0 = (S_h^0, E_h^0, I_1^0, I_2^0, R_h^0, S_a^0, E_a^0, I_a^0) = \left(\frac{\pi_h}{\mu_h}, 0, 0, 0, 0, \frac{\pi_a}{\mu_a}, 0, 0\right). \tag{19}
$$

3.4. Basic Reproduction Number

The effective reproduction number, denoted by $R_c = \rho(FV^{-1})$ in the model 1, where ρ represents the spectral radius of the next generation matrix FV^{-1}, is the number of new infections that Ascariasis-infected individuals cause when they interact with susceptible populations.

$$\mathcal{F} = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{\beta_h q(1-\epsilon)(1-\theta)\, S_h}{N_a} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\beta_a S_a}{N_h} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{20}$$

and

$$\mathcal{G} = \begin{bmatrix} k_1 & 0 & 0 & 0 & 0 \\ -\sigma_1 & k_2 & 0 & 0 & 0 \\ 0 & -\sigma_2\,(1-\tau_1) & k_3 & 0 & 0 \\ 0 & 0 & 0 & k_5 & 0 \\ 0 & 0 & 0 & -\sigma_a & \mu_a \end{bmatrix}. \tag{21}$$

where

$$k_1 = \sigma_1 + \mu_h, k_2 = \sigma_2 + \mu_h, k_3 = \delta_h + \mu_h + \tau_2, k_4 = \omega + \mu_h, k_5 = \sigma_a + \mu_a, \tag{22}$$

The inverse of (21) is given below

$$\mathcal{G}^{-1} = \begin{bmatrix} {k_1}^{-1} & 0 & 0 & 0 & 0 \\ \frac{\sigma_1}{k_2 k_1} & {k_2}^{-1} & 0 & 0 & 0 \\ -\frac{\sigma_2(-1+\tau_1)\sigma_1}{k_2 k_1 k_3} & -\frac{\sigma_2(-1+\tau_1)}{k_2 k_3} & {k_3}^{-1} & 0 & 0 \\ 0 & 0 & 0 & {k_5}^{-1} & 0 \\ 0 & 0 & 0 & \frac{\sigma_a}{k_5 \mu_a} & {\mu_a}^{-1} \end{bmatrix}, \tag{23}$$

then, the $\mathcal{F}\mathcal{G}^{-1}$ is given by

$$\mathcal{F}\mathcal{G}^{-1} = \begin{bmatrix} 0 & 0 & 0 & \frac{\beta_h q(1-\epsilon)(1-\theta)\, S_h \sigma_a}{N_a k_5 \mu_a} & \frac{\beta_h q(1-\epsilon)(1-\theta)\, S_h}{N_a \mu_a} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -\frac{\beta_a S_a \sigma_2(-1+\tau_1)\sigma_1}{N_h k_2 k_1 k_3} & -\frac{\beta_a S_a \sigma_2(-1+\tau_1)}{N_h k_2 k_3} & \frac{\beta_a S_a}{N_h k_3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{24}$$

Therefore, the eigenvalues are given by

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{\sqrt{N_h^0 k_2 k_1 k_3 N_a^0 k_5 \mu_a \beta_a S_a^0 \sigma_2 \sigma_1 \beta_h q(1-\theta)\, S_h^0 \sigma_a (1-\epsilon-\tau_1+\tau_1\epsilon)}}{N_h^0 k_2 k_1 k_3 N_a^0 k_5 \mu_a} \\ -\frac{\sqrt{N_h^0 k_2 k_1 k_3 N_a^0 k_5 \mu_a \beta_a S_a^0 \sigma_2 \sigma_1 \beta_h q(1-\theta)\, S_h^0 \sigma_a (1-\epsilon-\tau_1+\tau_1\epsilon)}}{N_h^0 k_2 k_1 k_3 N_a^0 k_5 \mu_a} \end{bmatrix}, \tag{25}$$

putting all the denominators of equation (25) under the roots, we obtain

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ \sqrt{\frac{N_h^0 k_2 k_1 k_3 N_a^0 k_5 \mu_a \beta_a S_a^0 \sigma_2 \sigma_1 \beta_h q(1-\theta)\, S_h^0 \sigma_a (1-\epsilon-\tau_1+\tau_1\epsilon)}{\left(N_h^0 k_2 k_1 k_3 N_a^0 k_5 \mu_a\right)^2}} \\ -\sqrt{\frac{N_h^0 k_2 k_1 k_3 N_a^0 k_5 \mu_a \beta_a S_a^0 \sigma_2 \sigma_1 \beta_h q(1-\theta)\, S_h^0 \sigma_a (1-\epsilon-\tau_1+\tau_1\epsilon)}{\left(N_h^0 k_2 k_1 k_3 N_a^0 k_5 \mu_a\right)^2}} \end{bmatrix}, \tag{26}$$

at disease-free equilibrium $S_h^0 \leq N_h^0$ and $S_a^0 \leq N_a^0$, now, after simplification, the spectral radius can be given by

$$\rho(\mathcal{F}\mathcal{G}^{-1}) = \sqrt{\frac{\beta_a \sigma_1 \sigma_2 \beta_h q(1-\theta)\sigma_a(1-\epsilon)(1-\tau_1)}{k_1 k_2 k_3 k_5 \mu_a}}. \tag{27}$$

Therefore, the basic reproduction number $\mathcal{R}_c$ of system (1) is given by the spectral radius of the matrix $\mathcal{F}\mathcal{G}^{-1}$,

$$\mathcal{R}_c = \sqrt{\frac{\beta_a \sigma_1 \sigma_2 \beta_h q(1-\theta)\sigma_a(1-\epsilon)(1-\tau_1)}{k_1 k_2 k_3 k_5 \mu_a}}. \tag{28}$$

3.5. Stability of Disease-Free Equilibrium Globally

Theorem 4. *For the model* (1)*, the disease-free equilibrium (DFE) is globally-asymptotically stable (GAS) in* $\mathcal{D}$ *when* $\mathcal{R}_c < 1$*; when* $\mathcal{R}_c > 1$*, it is unstable.*

Proof. For the DFE's GAS to work, $\mathcal{R}_c < 1$ [14] must satisfy the two axioms $[G_1]$ and $[G_2]$. In the form, we rewrite the system.

$$\begin{aligned} \frac{dX_a}{dt} &= F(X_a, X_b), \\ \frac{dX_b}{dt} &= G(X_a, X_b) : G(X_a, 0) = 0, \end{aligned} \tag{29}$$

where $X_b = (E_h^0, I_1^0, I_2^0, E_a^0, I_a^0)$ and $X_a = (S_h^0, R_h^0, S_a^0)$, where the uninfected population is represented by the elements of $X_a \in R_+^3$ and the infected population is represented by the elements of $X_b \in R_+^5$.
The DFE's current symbol is $M^0 = (X_a^*, 0)$,
where,
$X_a^* = (N_h^0/N_a^0, 0)$. Now let's discuss the first scenario. X_a^*'s globally asymptotic stability yields

$$\frac{dX_a}{dt} = F(X_1, 0) = \begin{bmatrix} \pi_h + \omega R_h - \mu_h S_h^0 \\ -(\mu_h + \omega) R_h^0 \\ \pi_a - \mu_a S_a^0 \end{bmatrix}. \tag{30}$$

A linear ODE solving gives

$$\begin{aligned} \frac{\pi_h + \omega R_h}{\mu_h} - \frac{\pi_h + \omega R_h}{\mu_h} e^{-\mu_h t} + S_h^0(0) e^{-\mu_h t} &= S_h^0(t), \\ -\frac{1}{(\omega + \mu_h)} e^{-(\omega+\mu_h)t} + R_h^0(0) e^{-(\omega+\mu_h)t} &= R_h^0(t), \\ \frac{\pi_h}{\mu_a} - \frac{\pi_a}{\mu_a} e^{-\mu_a t} + S_a^0(0) e^{-\mu_a t} &= S_a^0(t). \end{aligned}$$

Now, clearly from system (1) we have, $S_h^0(t) + R_h^0(t) \to N_h^0(t)$ and $S_a^0(t) \to N_a^0(t)$ as $t \to \infty$ regardless of the value of $S_h^0(t), R_h^0(t)$, and $S_a^0(t)$. Thus , $X_a^* = (N_h^0/N_a^0, 0)$ is stable asymptotically on a global scale.

Afterwards, regarding the second requirement, which is $\tilde{G}(X_a, X_a) = AX_b - G(X_a, X_b) \geq 0$

$$A = \begin{pmatrix} -k_1 & 0 & 0 & 0 & \frac{\beta_h q(1-\epsilon)(1-\theta)S_h}{N_a} \\ \sigma_1 & -k_2 & 0 & 0 & 0 \\ 0 & \sigma_2(1-\tau_1) & 0 & 0 & 0 \\ 0 & \frac{\beta_a \eta S_a}{N_h} & \frac{\beta_a S_a}{N_h} & -k_5 & 0 \\ 0 & 0 & 0 & \sigma_a & -\mu_a \end{pmatrix}. \tag{31}$$

As the matrix A has non-negative off-diagonal components, it is a Metziller matrix.

$$G(X_a, X_b) = \begin{pmatrix} \frac{\beta_h q(1-\epsilon)(1-\theta)I_a S_h}{N_a} - k_1 E_h, \\ \sigma_1 E_h - k_2 I_1, \\ \sigma_2(1-\tau_1)I_1 - k_3 I_2, \\ \frac{\beta_a (I_2 + \eta I_1) S_a}{N_h} - k_5 E_a, \\ \sigma_a E_a - \mu_a I_a \end{pmatrix}. \tag{32}$$

Then,

$$\tilde{G}(X_a, X_a) = AX_b - G(X_a, X_b) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

That is

$$\tilde{G}(X_a, X_b) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T.$$

Clearly, $\tilde{G}(X_a, X_b) = 0$. □

According to the epidemiological implications of the theorem 4, as long as the control reproduction number is less than one, the disease can be controlled in society. It is important to emphasize that one of the primary methods for reducing the number of reproductions is to control Ascariasis in society.

3.6. Endemic equilibrium point

Let

$$\Phi^{**} = (S_h^{**}, E_h^{**}, I_1^{**}, I_2^{**}, R_h^{**}, S_a^{**}, E_a^{**}, I_a^{**}) \tag{33}$$

be the endemic equilibrium, to achieve this, we let the right-hand side of equation (1) to zero and solve simultaneously at the endemic equilibrium, we have.

$$\begin{aligned}
S_h^{**} &= \frac{k_1 k_2 k_3 k_4 \pi_h}{(-\sigma_2 (k_3\tau_1 + \tau_2 (1-\tau_1)) \omega \sigma_1 + k_1 k_2 k_3 k_4) \lambda_h^{**} + k_3 k_4 k_2 k_1 \mu_h}, \\
E_h^{**} &= \frac{k_2 k_3 k_4 \pi_h \lambda_h^{**}}{(-\sigma_2 (k_3\tau_1 + \tau_2 (1-\tau_1)) \omega \sigma_1 + k_1 k_2 k_3 k_4) \lambda_h^{**} + k_3 k_4 k_2 k_1 \mu_h}, \\
I_1^{**} &= \frac{k_3 k_4 \sigma_1 \lambda_h^{**} \pi_h}{(-\sigma_2 (k_3\tau_1 + \tau_2 (1-\tau_1)) \omega \sigma_1 + k_1 k_2 k_3 k_4) \lambda_h^{**} + k_3 k_4 k_2 k_1 \mu_h}, \\
I_2^{**} &= \frac{\sigma_2 (1-\tau_1) k_4 \sigma_1 \pi_h \lambda_h^{**}}{(-\sigma_2 (k_3\tau_1 + \tau_2 (1-\tau_1)) \omega \sigma_1 + k_1 k_2 k_3 k_4) \lambda_h^{**} + k_3 k_4 k_2 k_1 \mu_h}, \\
R_h^{**} &= \frac{\sigma_2 (k_3\tau_1 + \tau_2 (1-\tau_1)) \sigma_1 \pi_h \lambda_h^{**}}{(-\sigma_2 (k_3\tau_1 + \tau_2 (1-\tau_1)) \omega \sigma_1 + k_1 k_2 k_3 k_4) \lambda_h^{**} + k_3 k_4 k_2 k_1 \mu_h}, \\
S_a^{**} &= \frac{\pi_a}{\lambda_a^{**} + \mu_a}, E_a^{**} = \frac{\pi_a \lambda_a^{**}}{k_5(\lambda_a^{**} + \mu_a)}, I_a^{**} = \frac{\sigma_a \pi_a \lambda_a^{**}}{\mu_a k_5 (\lambda_a^{**} + \mu_a)}
\end{aligned} \tag{34}$$

3.7. Stability of the Endemic Equilibrium Point Globally

Theorem 5. *In the model* (1)*, the endemic equilibrium (EE),* ϵ^**, is globally asymptotically stable (GAS) when* $\mathcal{R}_c > 1$ *and unstable when* $\mathcal{R}_c < 1$*.*

Proof. Let $\mathcal{V}$ be a Lyapunov function of the Goh-Volterra type, as shown below.

$$\begin{aligned}\mathcal{V} &= \left(S_h - S_h^{**} - S_h^{**} ln\frac{S_h}{S_h^{**}}\right) + \left(E_h - E_h^{**} - E_h^{**} ln\frac{E_h}{E_h^{**}}\right) + \frac{\sigma_1+\mu_h}{\sigma_1}\left(I_1 - I_1^{**} - I_1^{**} ln\frac{I_1}{I_1^{**}}\right)\\ &+ \frac{(\mu_h+\sigma_1)(\mu_h+\sigma_2)}{\sigma_1\sigma_2}\left(I_2 - I_2^{**} - I_2^{**} ln\frac{I_2}{I_2^{**}}\right)\\ &+ \frac{(\mu_h+\sigma_1)(\mu_h+\sigma_2)(\mu_h+\delta_h+\tau_2)}{\sigma_1\sigma_2\tau_2}\left(R_h - R_h^{**} - R_h^{**} ln\frac{R_h}{R_h^{**}}\right)\\ &+ \left(S_a - S_a^{**} - S_a^{**} ln\frac{S_a}{S_a^{**}}\right) + \left(E_a - E_a^{**} - E_a^{**} ln\frac{E_a}{E_a}\right) + \frac{(\mu_a+\sigma_a)}{\sigma_a}\left(I_a - I_a^{**} - I_a^{**} ln\frac{I_a}{I_a^{**}}\right),\end{aligned} \tag{35}$$

when we differentiate (35) in relation to time, we obtain

$$\begin{aligned}\dot{\mathcal{V}} &= \left(1 - \frac{S_h^{**}}{S_h}\right)\dot{S}_h + \left(1 - \frac{E_h^{**}}{E_h}\right)\dot{E}_h + \frac{\sigma_1+\mu_h}{\sigma_1}\left(1 - \frac{I_1^{**}}{I_1}\right)\dot{I}_1 + \frac{(\mu_h+\sigma_1)(\mu_h+\sigma_2)}{\sigma_1\sigma_2}\left(1 - \frac{I_2^{**}}{I_2}\right)\dot{I}_2\\ &+ \frac{(\mu_h+\sigma_1)(\mu_h+\sigma_2)(\mu_h+\delta_h+\tau_2)}{\sigma_1\sigma_2\tau_2}\left(1 - \frac{R_h^{**}}{R_h}\right)\dot{R}_h + \left(1 - \frac{S_a^{**}}{S_a}\right)\dot{S}_a + \left(1 - \frac{E_a^{**}}{E_a}\right)\dot{E}_a\\ &+ \frac{(\mu_a+\sigma_a)}{\sigma_a}\left(1 - \frac{I_a^{**}}{I_a}\right)\dot{I}_a,\end{aligned} \tag{36}$$

with

$$\begin{aligned}\pi_h &= \lambda_h^{**} S_h^{**} + \mu_h S_h^{**},\\ \pi_a &= \lambda_a^{**} S_a^{**} + \mu_a S_a^{**}.\end{aligned} \tag{37}$$

where

$$\begin{aligned}\lambda_h^{**} &= \frac{\beta_h q(1-\epsilon)(1-\theta)I_a^{**}}{N_a^{**}},\\ \lambda_a^{**} &= \frac{\beta_a(I_2^{**} + \eta I_1^{**}}{N_h^{**}}\end{aligned} \tag{38}$$

Substituting (1) into (36), we have

$$\begin{aligned}\dot{\mathcal{V}} &= \left(1 - \frac{S_h^{**}}{S_h}\right)(\lambda_h S_h^{**} + \mu_h S_h^{**} - \lambda_h S_h - \mu_h S_h)\\ &+ \left(1 - \frac{E_h^{**}}{E_h}\right)(\lambda_h S_h - (\sigma_1+\mu_h)E_h) + \frac{\sigma_1+\mu_h}{\sigma_1}\left(1 - \frac{I_1^{**}}{I_1}\right)(\sigma_h E_h - (\mu_h+\sigma_2)I_1)\\ &+ \frac{(\mu_h+\sigma_1)(\mu_h+\sigma_2)}{\sigma_1\sigma_2}\left(1 - \frac{I_2^{**}}{I_2}\right)(\sigma_2 I_1 - (\mu_h+\delta_h+\tau_2)I_2)\\ &+ \frac{(\mu_h+\sigma_1)(\mu_h+\sigma_2)(\mu_h+\delta_h+\tau_2)}{\sigma_1\sigma_2\tau_2}\left(1 - \frac{R_h^{**}}{R_h}\right)(\tau_2 I_2 - (\mu_h+\omega)R_h)\\ &+ \left(1 - \frac{S_a^{**}}{S_a}\right)(\lambda_a S_a^{**} + \mu_a S_a^{**} - \lambda_a S_a - \mu_a S_a) + \left(1 - \frac{E_a^{**}}{E_a}\right)(\lambda_a S_a - (\sigma_a+\mu_a)E_a)\\ &+ \frac{(\mu_a+\sigma_a)}{\sigma_a}\left(1 - \frac{I_a^{**}}{I_a}\right)(\sigma_a E_a - \mu_a I_a),\end{aligned} \tag{39}$$

with relations

$$\begin{aligned}(\sigma_1+\mu_h) &= \frac{\lambda_h S_h^{**}}{E_h^{**}},\\ (\sigma_2+\mu_h) &= \frac{\sigma_1 E_h^{**}}{I_1^{**}},\\ \mu_h &= \frac{\tau_2 I_2^{**}}{R_h^{**}},\\ (\sigma_a+\mu_a) &= \frac{\lambda_a S_a^{**}}{E_a^{**}},\\ \mu_a &= \frac{\sigma_a E_a^{**}}{I_a^{**}},\end{aligned} \tag{40}$$

substituting the relations in (40) into (39), we have

$$\begin{aligned}\dot{\mathcal{V}} \le{}& \mu_h S_h^{**}\left(2-\frac{S_h}{S_h^{**}}-\frac{S_h^{**}}{S_h}\right)+\lambda_h^{**}S_h^{**}\left(6-\frac{S_h^{**}}{S_h}-\frac{S_h E_h^{**}}{S_h^{**}E_h}-\frac{E_h I_1^{**}}{E_h^{**}I_1}-\frac{I_1 I_2^{**}}{I_1^{**}I_2}-\frac{R_h}{R_h^{**}}-\frac{I_2 R_h^{**}}{I_2^{**}R_h}\right)\\ &+\mu_a S_a^{**}\left(2-\frac{S_a}{S_a^{**}}-\frac{S_a^{**}}{S_a}\right)+\lambda_a S_a^{**}\left(4-\frac{S_a^{**}}{S_a}-\frac{S_a E_a^{**}}{S_a^{**}E_a}-\frac{I_a}{I_a^{**}}-\frac{E_a I_a^{**}}{E_a^{**}I_a}\right).\end{aligned} \tag{41}$$

Then, using the relationship between the arithmetic and geometric means, we have

$$\begin{aligned}&\left(2-\frac{S_h}{S_h^{**}}-\frac{S_h^{**}}{S_h}\right)\le 0,\left(6-\frac{S_h^{**}}{S_h}-\frac{S_h E_h^{**}}{S_h^{**}E_h}-\frac{E_h I_1^{**}}{E_h^{**}I_1}-\frac{I_1 I_2^{**}}{I_1^{**}I_2}-\frac{R_h}{R_h^{**}}-\frac{I_2 R_h^{**}}{I_2^{**}R_h}\right)\le 0,\\ &\left(2-\frac{S_a}{S_a^{**}}-\frac{S_a^{**}}{S_a}\right)\le 0,\left(4-\frac{S_a^{**}}{S_a}-\frac{S_a E_a^{**}}{S_a^{**}E_a}-\frac{I_a}{I_a^{**}}-\frac{E_a I_a^{**}}{E_a^{**}I_a}\right)\le 0.\end{aligned} \tag{42}$$

With the following criteria, we have $\dot{\mathcal{V}}\le 0$: $\omega=0$ and $\mathcal{R}_c>1$. Since all of the variables in the model that are involved, including $S_h, E_h, I_1, I_2, R_h, S_a, E_a, I_a$, are at steady state, this can be replaced into the concerned variable of (1) to create

$$\lim_{t\to\infty}\left(S_h(t),E_h(t),I_1(t),I_2(t),R_h(t),S_a(t),E_a(t),I_a(t)\right)\to\left(S_h,E_h,I_1,I_2,R_h,S_a,E_a,I_a\right). \tag{43}$$

Accordingly, using Lassalle's invariant principle, the endemic equilibrium point is globally asymptotically stable (GAS) [15–19, 27–32]. □

4. Numerical simulations

This section displays the state of the model as simulated numerically using the parameter values in Table 2. An analysis of numerical simulations has been conducted to comprehend the transmission dynamics of the model (1) fully, we used the ODE45 solver in MATLAB 2014. Time-series diagrams were used to illustrate the behaviour of the compartments and how essential factors influence state variables. All the parameter values used are as given in Table 2.

Table 2. Ranges and baseline values of parameters of model (1).

Parameter	Ranges (Baseline)	Unit	Reference
π_h, π_a	$100, 2000$	$year^{-1}$	[21, 22, 25, 26]
μ_h, μ_a	$0.0004, 0.020$	$year^{-1}$	[20]
β_h, β_a	$0.0005, 0.800$	$year^{-1}$	[20]
$\sigma_1, \sigma_2, \sigma_a$	$0.0004, 0.020$	$year^{-1}$	[20]
τ_1, τ_2	$(0,1), (0,1)$	$year^{-1}$	Control parameter
q	0.05	$year^{-1}$	Assume
η	0.17	$year^{-1}$	Assumed
ω	0.034	$year^{-1}$	Assume
ϵ	$(0,1)$	$year^{-1}$	Control parameter
δ_h	0.02	$year^{-1}$	[24]

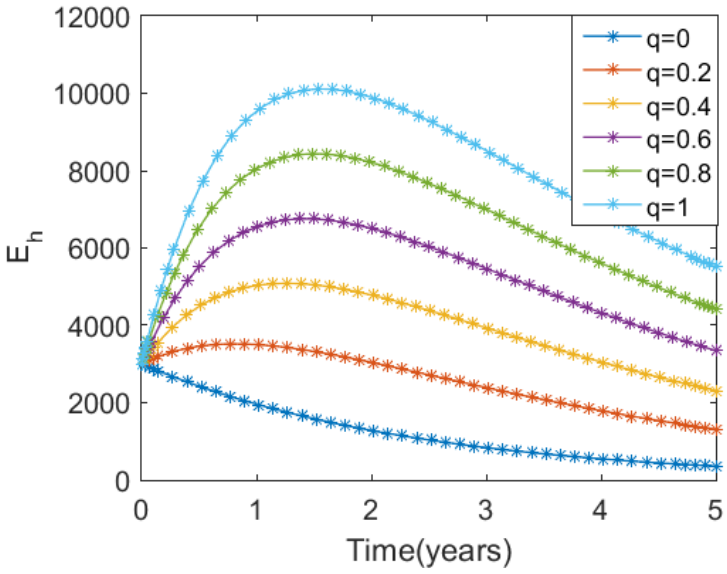


Figure 2. Quantity of not done animal flesh consumed influence on the Exposed humans compartment

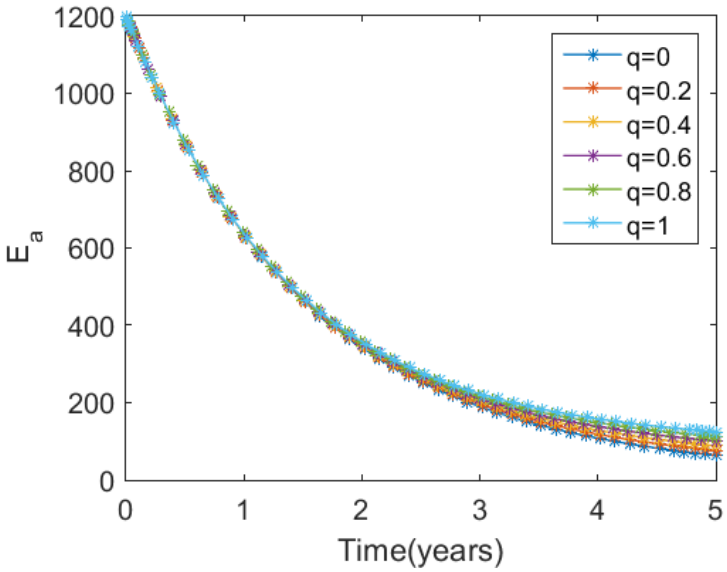


Figure 3. Quantity of not done animal flesh consumed influence on the Exposed animals compartment

5. Examining simulated figures

Figure 2 describes how the quantity of not done animal flesh consumed affects the dynamics of exposed humans. As consumption increases, the population of exposed humans also increases; this indicates that the consumption of improperly prepared animal flesh is directly affecting the dynamics of exposed humans. Figure 3 describes how the quantity of not done animals' flesh consumed affects the dynamics of exposed animals. As consumption increases, the population of exposed animals also increases slightly; this suggests that the consumption of not yet exposed animals' flesh has no direct effect on the dynamics of exposed animals. Figure 4 describes how the quantity of not done animal flesh consumed affects the dynamics of the Light-infected humans compartment. As the consumption increases, the population of the Light-infected humans compartment also increases; this shows that the consumption of not done animal flesh is directly

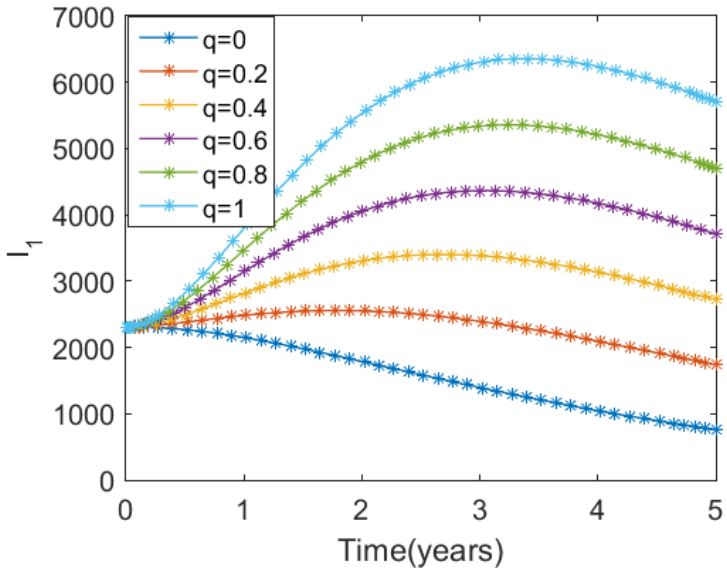


Figure 4. Quantity of not done animal flesh consumed influence on the Light infected humans compartment

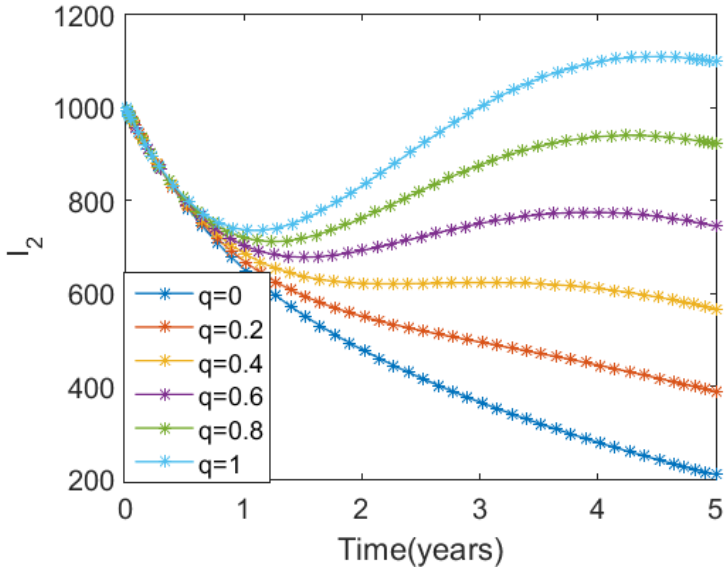


Figure 5. Quantity of not done animal flesh consumed influence on the Heavy infectious human compartment

affecting the dynamics of the Light-infected humans compartment. Figure 5 describes how the quantity of not done animal flesh consumed affects the dynamics of the heavy infectious human compartment. As the consumption increases, the population of the heavy infectious humans compartment also increases; this shows that the consumption of not done animal flesh is directly affecting the dynamics of the heavy infectious humans compartment. Figure 6 describes how the quantity of not done animal flesh consumed affects the dynamics of the contagious animal compartment. As the consumption increases, the population of the infectious animal compartment remains unchanged. This suggests that the consumption of improperly prepared animal flesh has no connection to the dynamics of the contagious animal compartment.

Figure 7 illustrates how early treatment impacts the dynamics of exposed individuals. As the early treatment increases, the population of exposed humans decreases slightly; this indicates that the early treatment has a minimal effect on the dynamics of exposed humans. Figure 8 describes

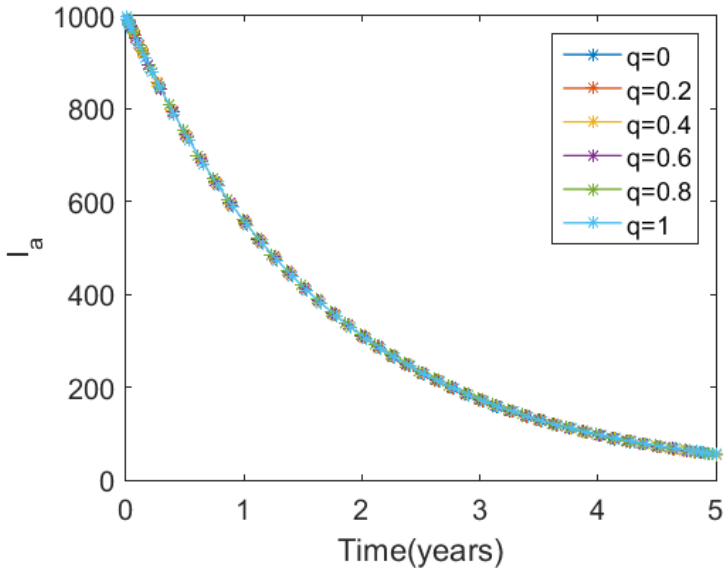


Figure 6. Quantity of not done animal flesh consumed influence on the infectious animal compartment

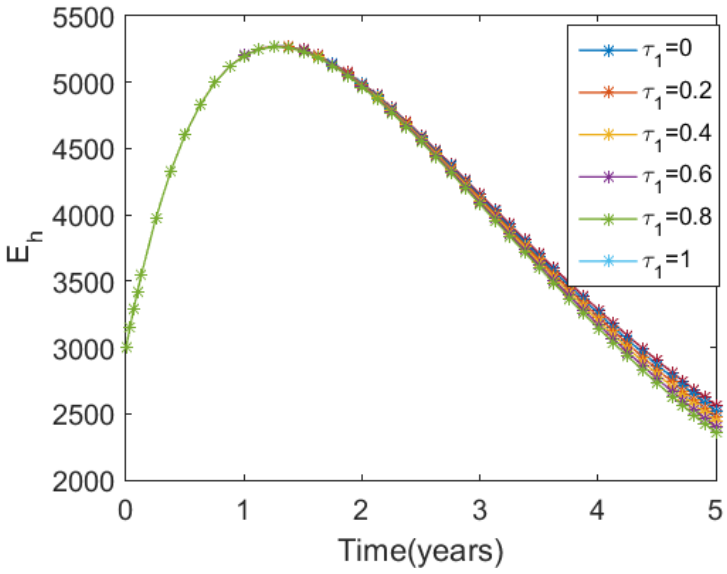


Figure 7. Early treatment influence on the Exposed humans compartment

how the quantity of early treatment affects the dynamics of exposed animals. As the early treatment increases, the population of exposed animals also decreases slightly; this indicates that the early treatment has a minimal effect on the dynamics of exposed animals. Figure 9 describes how early treatment affects the dynamics of the Light-infected humans compartment. As the early treatment increases, the population of the Light-infected humans compartment decreases slightly; this indicates that the early treatment has a minimal effect on the dynamics of the Light-infected humans compartment. Figure 10 describes how early treatment affects the dynamics of the heavy infectious human compartment. As the early treatment increases, the population of the heavy infectious humans compartment decreases drastically, which shows that the early treatment is directly affecting the dynamics of the heavy infectious humans compartment. Figure 11 illustrates how early treatment impacts the dynamics of the contagious animal compartment. As the early treatment increases the population of the infectious animals compartment, it remains

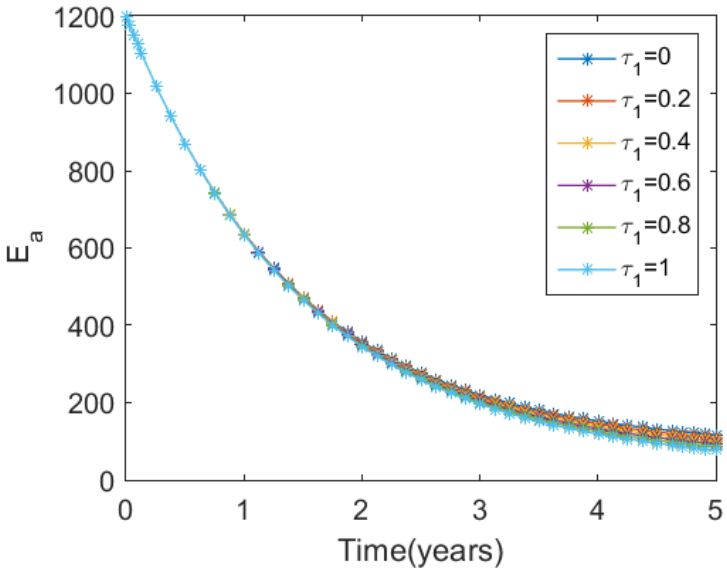


Figure 8. Early treatment influence on Exposed animals compartment

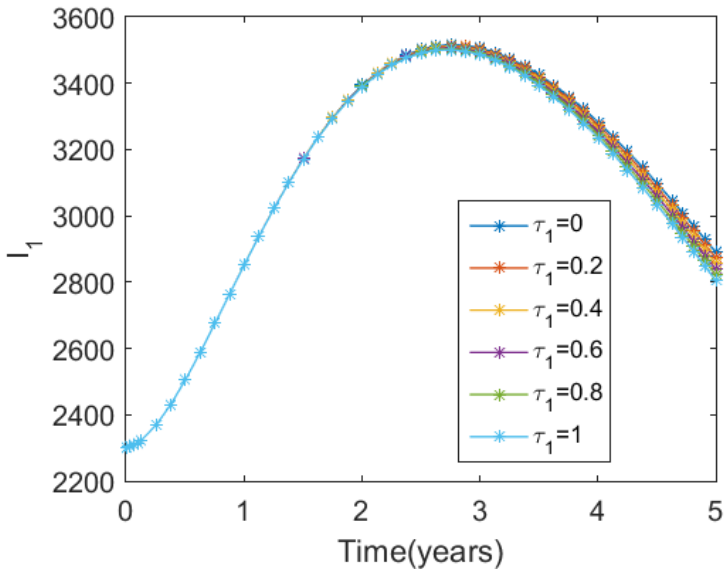


Figure 9. Early treatment influence on Light-infected human compartment

unchanged, which shows that the early treatment has no connection with the dynamics of the contagious animals compartment.

Figure 12 illustrates how the treatment of the heavily infected human compartment impacts the dynamics of exposed individuals. As the treatment of the heavily infectious human compartment increases, the population of exposed humans decreases slightly, indicating that the treatment of the heavily infectious human compartment has a minimal effect on the dynamics of exposed humans. Figure 13 illustrates how the treatment of the heavily infected human compartment affects the dynamics of exposed animals. As the treatment of heavily infected human compartments increases, the population of exposed animals also decreases slightly, indicating that the treatment of heavily infected human compartments has a minimal effect on the dynamics of exposed animals. Figure 14 describes how treatment of the Heavy infectious humans compartment affects the dynamics of the Light infected humans compartment, as the treatment of the Heavy infectious humans compartment increases, the population of Light infected humans compartment decreases a little, this shows that the treatment of the Heavy infectious humans compartment has

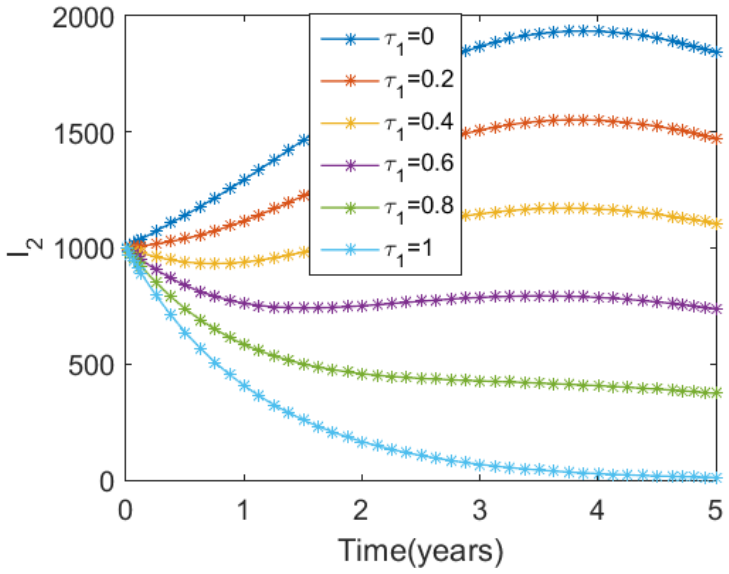


Figure 10. Early treatment influence on the heavily infectious human compartment

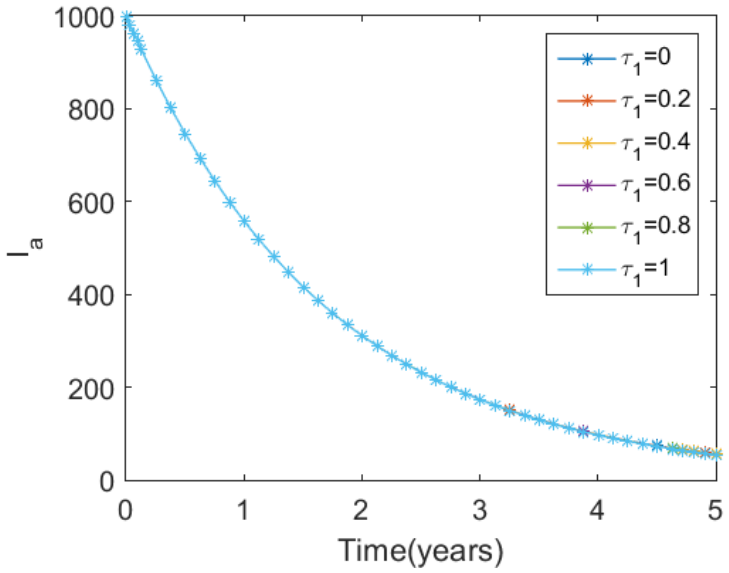


Figure 11. Early treatment influence on infectious animal compartments

a little effect on the dynamics of Light infected humans compartment. Figure 15 describes how the treatment of the Heavy infectious humans compartment affects the dynamics of the heavy infectious humans compartment. As the treatment of the heavily infectious human compartment increases, the population of the heavily infectious human compartment decreases drastically. This indicates that early treatment has a direct impact on the dynamics of the heavily infected human compartment. Figure 16 describes how the treatment of the heavily infected human compartment affects the dynamics of the infectious animal compartment. As the treatment of the heavily infected human compartment increases, the population of the infectious animals compartment remains unchanged, indicating that the treatment of the heavily infected human compartment has no connection with the dynamics of the infectious animals compartment.

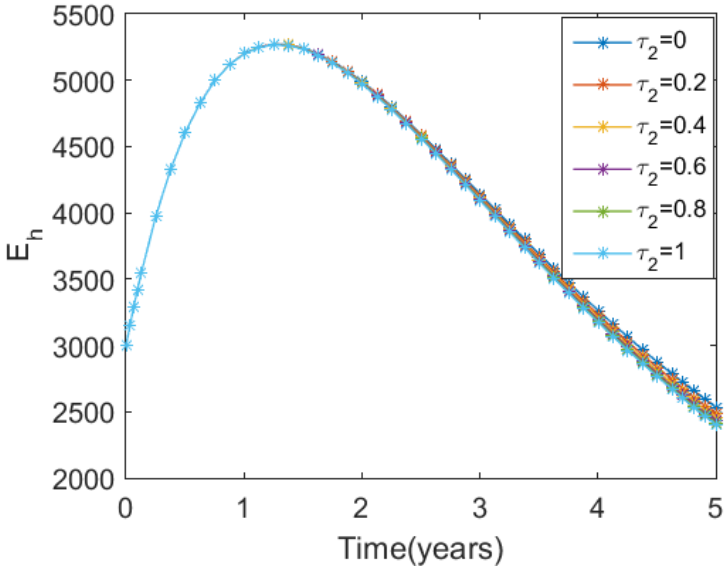


Figure 12. Treatment rate of heavily infectious humans influence on the Exposed humans compartment

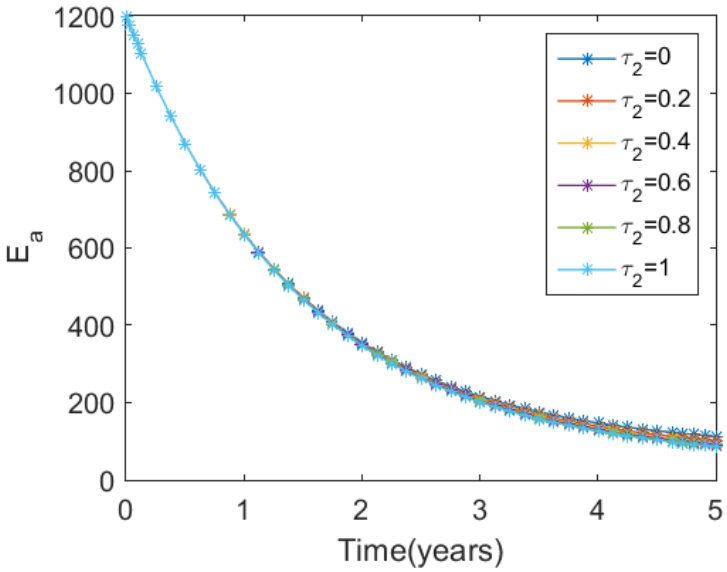


Figure 13. Treatment rate of heavily infectious humans influence on the Exposed animals compartment

6. Conclusion

This work considers both the human and animal populations. The human population has been subdivided into five compartments, including susceptible humans, exposed humans, lightly infected humans, heavily infected humans, and recovered humans. The animal population has been subdivided into three groups: susceptible animals, exposed animals, and infectious animals. The formulated models have been analyzed. To begin the analysis, we verified the existence and uniqueness of the proposed model, followed by the boundedness and positivity of the solutions. Next, we computed the practical and basic reproduction numbers, and we also checked the local stability of the disease-free equilibrium using a Jacobian matrix. This analysis shows that the disease-free equilibrium is locally asymptotically stable if the effective reproduction number is

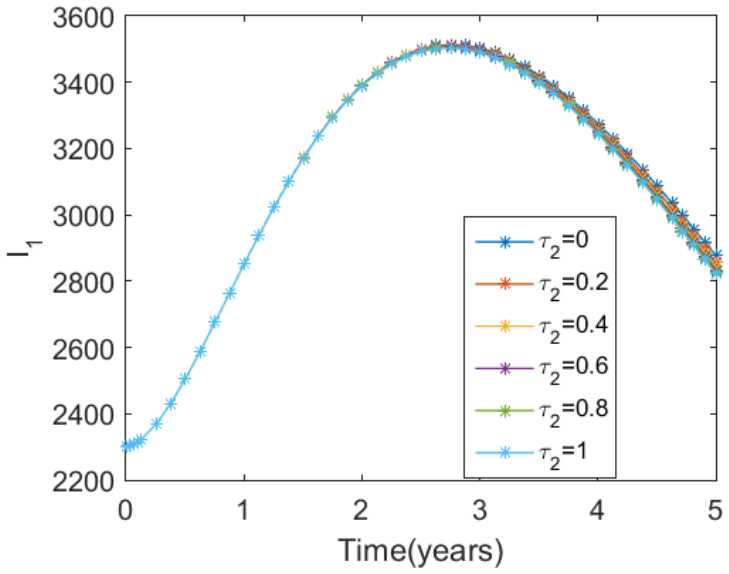


Figure 14. Treatment rate of heavily infectious humans influence on the Light-infected humans compartment

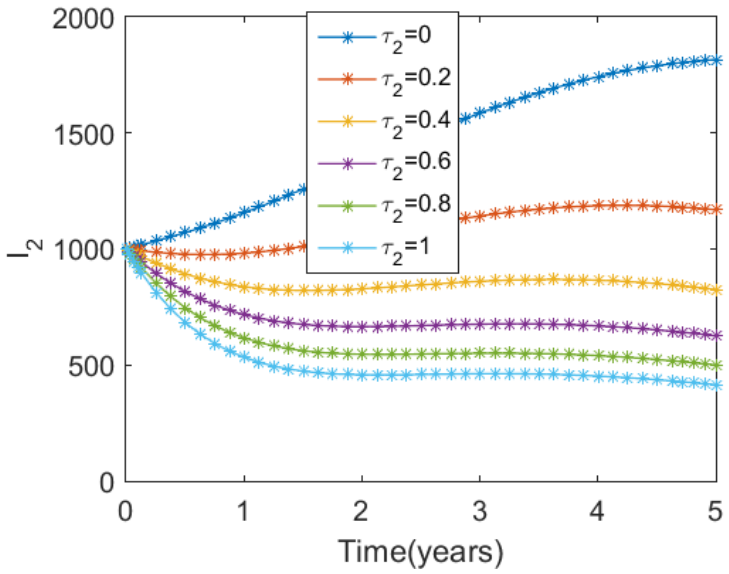


Figure 15. Treatment rate of heavily infectious humans influence on the Heavy infectious humans compartment

less than unity. The global asymptotic stability of the disease-free equilibrium has been investigated using the comparison method, which demonstrates that the disease-free equilibrium is globally asymptotically stable if the effective reproduction number is less than one. The global asymptotic stability of the endemic equilibrium has been studied using the Lyapunov function of Go-Volterra type, which shows that the endemic equilibrium point is globally asymptotically stable if the effective reproduction number is greater than unity. Numerical simulation shows that early treatment plays a vital role in controlling the dynamics of Ascariasis, as well. It has been observed that consuming raw animal flesh can increase or exacerbate diseases in society. This means that reducing the consumption of raw animal flesh can help control the Ascariasis disease in society. It is therefore recommended that early treatment should be prioritized, and eating the flesh of uncooked animals should be reduced in order to control the disease in society.

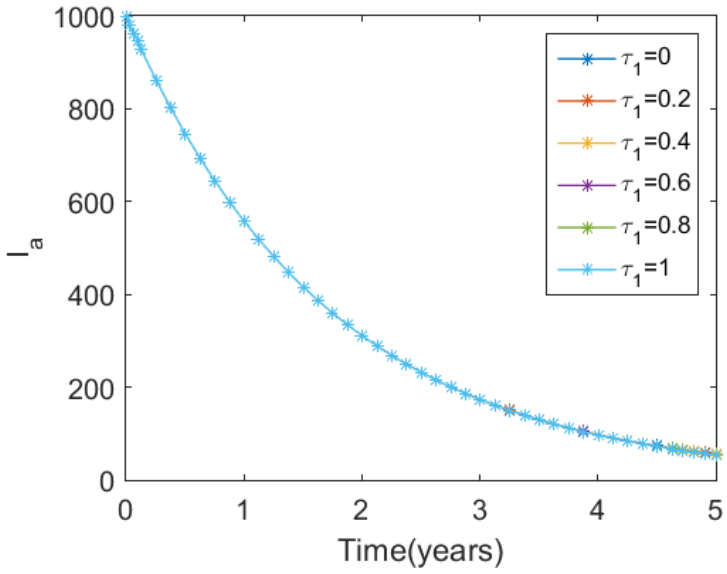


Figure 16. Treatment rate of heavily infectious humans influence on the infectious animals compartment

Acknowledgments

None.

Funding

None.

Conflict of interest

There is no conflict of interest to disclose.

Author contributions

James Andrawus: Methodology, Conceptualization, Formal analysis, Writing - Original draft. **Ibrahim Ibrahim Muhammad:** Methodology, Conceptualization, Formal analysis, Writing - Original draft. **Felix Yakubu Eguda:** Methodology, Conceptualization, Formal analysis, Writing - Original draft. **Abdullahi Yusuf:** Validation, Investigation, Formal analysis, Supervision

Declaration of using AI tools

The authors declare that they have not used any type of generative artificial intelligence for the writing of this manuscript, nor for the creation of images, graphics, tables, or their corresponding captions.

References

[1] Centers for Disease Control and Prevention (CDC), Ascariasis. Retrieved from `https://www.cdc.gov/parasites/ascariasis` 2023.

[2] F. E. G. Cox, "History of human parasitology," Clin. Microbiol. Rev., vol. 15, no. 4, pp. 595–612, Oct. 2002. `https://doi.org/10.1128/CMR.15.4.595-612.2002`

[3] J. Bethony, S. Brooker, M. Albonico, S. M. Geiger, A. Loukas, D. Diemert, et al., "Soil-transmitted helminth infections: Ascariasis, trichuriasis, and hookworm," Lancet, vol. 367, no. 9521, pp. 1521–1532, May 6 2006. `https://doi.org/10.1016/S0140-6736(06)68653-4`

[4] S. Brooker, A. C. A. Clements, and D. A. P. Bundy, "Global epidemiology, ecology and control of soil-transmitted helminth infections," Adv. Parasitol., vol. 62, pp. 221–261, 2006. https://doi.org/10.1016/S0065-308X(05)62007-6

[5] P. J. Hotez, D. H. Molyneux, A. Fenwick, J. Kumaresan, S. E. Sachs, J. D. Sachs, et al., "Control of neglected tropical diseases," N. Engl. J. Med., vol. 357, no. 10, pp. 1018–1027, Sep. 6 2007., https://doi.org/10.1056/NEJMra064142

[6] N. R. de Silva, S. Brooker, P. J. Hotez, A. Montresor, D. Engels, and L. Savioli, "Soil-transmitted helminth infections: Updating the global picture," Trends Parasitol., vol. 19, no. 12, pp. 547–551, Dec. 2003., https://doi.org/10.1016/j.pt.2003.10.002

[7] Centers for Disease Control and Prevention (CDC), Ascariasis, 2022. https://www.cdc.gov/parasites/ascariasis/index.html

[8] World Health Organisation (WHO), "Soil-transmitted helminth infections", 2003. https://www.who.int/news-room/fact-sheets/detail/soil-transmitted-helminth-infections

[9] P. M. Jourdan, P. H. L. Lamberton, A. Fenwick, and D. G. Addiss, "Soil-transmitted helminth infections," Lancet, vol. 391, no. 10117, pp. 252–265, Jan. 20 2018. https://doi.org/10.1016/S0140-6736(17)31930-X

[10] P. R. Hollingsworth, Jr., A. M. Simons, J. A. Fordyce, and C. D. Hulsey, "Explosive diversification following a benthic to pelagic shift in freshwater fishes," BMC Evol. Biol., vol. 13, no. 272, p. 272, Dec. 17 2013. https://doi.org/10.1186/1471-2148-13-272

[11] M. Betson, P. Nejsum, R. P. Bendall, R. M. Deb, and J. R. Stothard, "Molecular epidemiology of ascariasis: A global perspective on the transmission dynamics of Ascaris in people and pigs," J. Infect. Dis., vol. 210, no. 6, pp. 932–941, Sep. 15 2014. https://doi.org/10.1093/infdis/jiu193

[12] K. Parvatham, L. Veerakumari, and K. Sankarasubramanian, "A mathematical approach to the carbohydrate catabolism in Ascaris lumbricoides," Indian J. Sci. Technol., vol. 5, no. 5, pp. 1–5, 2012. https://doi.org/10.17485/ijst/2012/v5i5.2

[13] A. C. Fowler and T. D. Hollingsworth, "The dynamics of Ascaris lumbricoides infections," Bull. Math. Biol., vol. 78, no. 4, pp. 815–833, Apr. 2016. https://doi.org/10.1007/s11538-016-0164-2

[14] P. van den Driessche and J. Watmough, "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission," Math. Biosci., vol. 180, no. 1-2, pp. 29–48, Nov-Dec 2002. https://doi.org/10.1016/S0025-5564(02)00108-6

[15] J. P. Lasalle, "The stability of dynamical systems," in Regional conference series in applied mathematics. Philadelphia: SIAM, 1976.

[16] J. Andrawus, A. Iliyasu Muhammad, B. Akawu Denue, H. Abdul, A. Yusuf, and S. Salahshour, "Unraveling the importance of early awareness strategy on the dynamics of drug addiction using mathematical modeling approach," Chaos, vol. 34, no. 8, p. 083112, Aug. 1 2024. https://doi.org/10.1063/5.0203892

[17] J. Andrawus, Y. U. Ahmad, A. A. Andrew, A. Yusuf, S. Qureshi, B. A. Denue, et al., "Impact of surveillance in human-to-human transmission of monkeypox virus," Eur. Phys. J. Spec. Top., vol. 234, no. 3, pp. 483–514, 2024. https://doi.org/10.1140/epjs/s11734-024-01346-5

[18] J. Andrawus, K. G. Ibrahim, I. Abdullahi, A. Abubakar, and S. I. Maiwa, Mathematical modelling on drug addiction with awareness control, vol. 5. 2024, pp. 107–122.

[19] G. Djatcha Yaleu, S. Bowong, E. Houpa Danga, and J. Kurths, "Mathematical analysis of the dynamical transmission of Neisseria meningitidis serogroup," Int. J. Comput. Math., vol. 94, no. 12, pp. 2409–2434, 2017. https://doi.org/10.1080/00207160.2017.1283411

[20] E. O. Alzahrani and M. A. Khan, "Modeling the dynamics of hepatitis E with optimal control," Chaos Solitons Fractals, vol. 116, pp. 287–301, 2018. https://doi.org/10.1016/j.chaos.2018.09.033

[21] M. Sholikah and C. Alfiniyah, "Miswanto, "Stability analysis and optimal control of mathematical model for the spread of hepatitis E"," Communications in Mathematical Biology and Neuroscience, vol. 2020, 2020.

[22] M. Uema, K. Yonemitsu, Y. Sasaki, and H. Asakura, "Detection of hepatitis E virus RNA from pig bile collected at a slaughterhouse in Japan," AIMS Microbiol., vol. 8, no. 4, pp. 566–574, Dec. 26 2022. https://doi.org/10.3934/microbiol.2022036

[23] P. Li, Y. Ji, Y. Li, Z. Ma, and Q. Pan, "Estimating the global prevalence of hepatitis E virus in swine and pork products," One Health, vol. 14, p. 100362, Dec. 14 2021. https://doi.org/10.1016/j.onehlt.2021.100362

[24] S. Osman, B. S. Lassong, M. Dasumani, E. Y. Boateng, W. M. Onsongo, B. Diallo, et al., "Modeling the transmission routes of hepatitis E Virus as a zoonotic disease using fractional-order derivative," J. Appl. Math., vol. 2024, no. 1, p. 5168873, 2024. https://doi.org/10.1155/2024/5168873

[25] E. Uçar, N. Özdemir, and E. Altun, "Qualitative analysis and numerical simulations of new model describing cancer," J. Comput. Appl. Math., vol. 422, no. 114899, p. 114899, 2023. https://doi.org/10.1016/j.cam.2022.114899

[26] S. Hirsch, D. Ölz, and C. Schmeiser, "Existence and uniqueness of solutions for a model of non-sarcomeric actomyosin bundles," Discrete Contin. Dyn. Syst. Ser. A, vol. 36, no. 9, pp. 4945–4962, 2016. https://doi.org/10.3934/dcds.2016014

[27] J. Andrawus, J. Y. Musa, S. Babuba, A. Yusuf, S. Qureshi, U. T. Mustapha, et al., "Modeling the dynamics of pertussis to assess the influence of timely awareness with optimal control analysis," Journal of the Nigerian Society of Physical Sciences, vol. 7, no. 4, p. 2732, 2025. https://doi.org/10.46481/jnsps.2025.2732

[28] J. Andrawusa, K. I. Omotoso, A. A. Andrew, F. Y. Eguda, S. Babuba, and K. G. Ibrahim, "Mathematical model analysis on the significance of surveillance and awareness on the transmission dynamics of diphtheria," Journal of the Nigerian Society of Physical Sciences, vol. 7, no. 4, p. 2618, 2025. https://doi.org/10.46481/jnsps.2025.2618

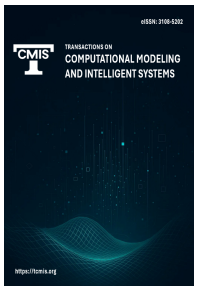
[29] K. Bitrus, F. A. Abdullah, M. O. Adewole, J. Andrawus, S. Qureshi, A. Soomro, et al., "Uncovering the influence of public awareness on dengue population dynamics: A mathematical model approach," Bol. Soc. Mat. Mex., vol. 31, no. 2, p. 86, 2025. https://doi.org/10.1007/s40590-025-00778-w

[30] U. T. Mustapha, A. I. Muhammad, A. Yusuf, N. Althobaiti, A. I. Aliyu, and J. Andrawus, "Mathematical dynamics of meningococcal meningitis: Examining carrier diagnosis and prophylaxis treatment," Int. J. Appl. Comput. Math., vol. 11, no. 2, p. 77, 2025. https://doi.org/10.1007/s40819-025-01890-1

[31] K. G. Ibrahim, J. Andrawus, A. Abubakar, A. Yusuf, S. I. Maiwa, K. Bitrus, et al., "Mathematical analysis of chickenpox population dynamics unveiling the impact of booster in enhancing recovery of infected individuals", Modeling Earth Systems and Environment, vol. 11, artcile 46, 2025. https://doi.org/10.1007/s40808-024-02219-5

[32] J. Andrawus, A. Abubakar, A. Yusuf, A. A. Andrew, B. Uzun, and S. Salahshour, "Impact of public awareness on haemo-lyphatic and meningo-encepphalitic stage of sleeping sickness using mathematical model approach," Eur. Phys. J. Spec. Top., vol. 234, no. 8, pp. 2005–2023, 2024. https://doi.org/10.1140/epjs/s11734-024-01417-7

All open access articles published in Transactions on Computational Modeling and Intelligent Systems (http://tcmis.org) are distributed under the terms of the CC BY-NC 4.0 license (Creative Commons Attribution Non-Commercial 4.0 International Public License as currently displayed at http://creativecommons.org/licenses/by-nc/4.0/legalcode) which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original work is properly cited.



Transactions on
Computational Modeling
and Intelligent Systems

Volume 2, January 2026, Article 10028
https://doi.org/10.65112.tcmis.10028

eISSN: 3108-5202
www.tcmis.org

REVIEW ARTICLE

A selective review of modern stochastic modeling: SDE/SPDE numerics, data-driven identification, and generative methods with applications in biomathematics

Yassine Sabbar[1] and Kottakkaran Sooppy Nisar[2,*]

[1] *IMIA Laboratory, T-IDMS, Department of Mathematics, FST Errachidia, Moulay Ismail University of Meknes, P.O. Box 509, 52000, Boutalamine, Morocco.*
[2] *Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia.*
[*] *Corresponding Author. Email: n.sooppy@psau.edu.sa (K. S. Nisar)*

Article Information

Received: 8 December 2025
Accepted: 18 December 2025
Published 7 January 2026

AMS Classification:
60H10, 60H30, 60J74, 92D30

Abstract

This review maps 2020-2025 developments in stochastic modeling, highlighting non-standard approaches and their applications to biology and epidemiology. It brings together four strands: (1) core models for systems that evolve with randomness; (2) learning key parts of those models directly from data; (3) methods that can generate realistic synthetic data in continuous time; and (4) numerical techniques that keep simulations stable, accurate, and faithful over long runs. The objective is practical: help researchers quickly see what is new, how the pieces fit together, and where important gaps remain. We summarize tools for estimating changing infection or reaction rates under noisy and incomplete observations, modeling spatial spread, accounting for sudden jumps and heavy tails, and reporting uncertainty in a way that is useful for decisions. We also highlight open problems that deserve near-term attention: separating true dynamics from noise when data are irregular; learning spatial dynamics under random influences with guarantees of stability; aligning training with the numerical method used in applications; preserving positivity and conservation in all simulations; reducing cost while controlling error for large studies; estimating rare but important events; and adopting clear, comparable reporting standards. By organizing the field around these aims, the review offers a concise guide to current methods, their practical use, and the most promising directions for future work in biology and epidemiology.

Keywords: Stochastic modeling, stochastic differential equations, stochastic partial differential equations, neural stochastic differential equations, operator learning, numerical methods

Nomenclatures

Abbrev.	Meaning	Abbrev.	Meaning
ABC	Approximate Bayesian Computation	MLE	Maximum Likelihood Estimation
CI	Confidence Interval	MLMC	Multilevel Monte Carlo
CRPS	Continuous Ranked Probability Score	NPE	Neural Posterior Estimation
DSM	Denoising Score Matching	ODE	Ordinary Differential Equation
EnKF	Ensemble Kalman Filter	PDE	Partial Differential Equation
ESS	Effective Sample Size	PINO	Physics-Informed Neural Operator
FD	Finite Difference	PINN	Physics-Informed Neural Network
IMEX	Implicit-Explicit	PSD	Positive Semidefinite
MCMC	Markov Chain Monte Carlo	SDE	Stochastic Differential Equation
SG-MCMC	Stochastic-Gradient Markov Chain Monte Carlo	SPDE	Stochastic Partial Differential Equation
SIR	Susceptible-Infected-Recovered	SMC	Sequential Monte Carlo
SIS	Susceptible-Infected-Susceptible	SRK	Stochastic Runge-Kutta
UQ	Uncertainty Quantification		

1. Introduction

Stochastic modeling offers a mathematically rigorous and physically interpretable framework for analyzing dynamical systems whose evolution is influenced by intrinsic randomness [1], unresolved multiscale interactions [2], or incomplete specification of the governing mechanisms [3]. In contrast to deterministic formulations (where the system trajectory is uniquely determined by a prescribed set of ordinary or partial differential equations [4]) stochastic models incorporate random variables or stochastic processes directly into the governing equations [5]. This probabilistic formulation enables explicit quantification of epistemic and aleatory uncertainty, reproduces variability across independent realizations [6], and captures non-deterministic phenomena, such as intermittency, heavy-tailed fluctuations, or regime switching, that lie beyond the representational capacity of purely deterministic models [7].

For finite-dimensional systems, the canonical representation is the stochastic differential equation (SDE):

$$\mathrm{d}X_t = f(X_t, t)\,\mathrm{d}t + G(X_t, t)\,\mathrm{d}W_t, \tag{1}$$

where $X_t \in \mathbb{R}^d$ denotes the state vector, $f : \mathbb{R}^d \times [0, \infty) \to \mathbb{R}^d$ is the drift function describing the deterministic dynamics, $G : \mathbb{R}^d \times [0, \infty) \to \mathbb{R}^{d\times m}$ is the diffusion coefficient matrix, and W_t is an m-dimensional standard Wiener process modeling temporally continuous Gaussian perturbations [8]. The drift term f typically encodes system-specific physical laws or conservation principles, whereas the diffusion term G parameterizes the intensity, anisotropy, and possible state-dependence of the noise [9]. When the state space is infinite-dimensional, one arrives at stochastic partial differential equations (SPDEs), where the solution $X(t, \xi)$ depends on both temporal (t) and spatial (ξ) variables, and the stochastic forcing may act locally or globally in space [10]. SPDEs arise naturally in the modeling of spatially distributed systems subject to random forcing, such as turbulent flows [11], population dispersal [12], or reaction-diffusion processes with environmental fluctuations [13].

The classical Itô calculus framework has been generalized to accommodate diverse noise structures [14], including: (i) *Jump-diffusion processes* [15], where compensated Poisson random measures introduce finite-activity discontinuities representing abrupt or rare events [16]; (ii) *Lévy-driven systems* [17], which allow for heavy-tailed increments and infinite-activity jump processes [18], capturing burst-like dynamics and extreme variability [19]; and (iii) *Fractional Brownian motion* and *Volterra-type kernels* [20, 21], which encode long-range temporal dependence, self-similarity, and memory effects [22]. In many applications, the diffusion term $G(X_t, t)$ is itself state-dependent, creating a multiplicative noise structure and introducing bidirectional coupling between the system dynamics and the stochastic perturbations [23].

From a statistical perspective, stochastic models serve both as *generative mechanisms* and as *inferential frameworks* [24, 25]. As generative models, they produce synthetic sample paths that preserve empirical distributional properties [26], temporal correlation structures [27], and

extreme-event statistics observed in experimental or observational datasets [28]. As inferential frameworks, they provide a principled foundation for recovering latent parameters, hidden states [29], and entire posterior distributions from noisy, incomplete, or irregularly sampled measurements [30].

Parameter estimation for SDEs and SPDEs is supported by a broad methodological arsenal, including maximum likelihood estimation [31], the generalized method of moments [32], and Bayesian inference frameworks [33] that yield full posterior distributions and enable comprehensive uncertainty quantification [34]. Computational advances now permit the application of Markov chain Monte Carlo (MCMC) methods (such as Hamiltonian Monte Carlo and particle MCMC [35]) to efficiently explore high-dimensional posterior spaces [36]. For online estimation in systems with streaming data, sequential Monte Carlo methods, such as the particle filter and the ensemble Kalman filter, have proven indispensable [37]. In scenarios where the likelihood function is analytically intractable, likelihood-free inference techniques, including approximate Bayesian computation and synthetic likelihood methods [38], offer practical alternatives by relying on forward simulation and carefully chosen summary statistics in place of explicit likelihood evaluation [39].

Recent theoretical advances in high-frequency sampling asymptotics have yielded consistent and asymptotically efficient estimators for both drift and diffusion coefficients under substantially weaker smoothness and regularity conditions than those previously required [40]. These results hold even under model misspecification [41], endogenous observation noise [42], or irregular sampling schedules, thereby extending their applicability to empirical datasets where idealized sampling assumptions are violated [43]. In the context of SPDEs, spectral estimation methods (combined with finite-dimensional Galerkin projections [44]) enable accurate recovery of dynamical modes from partial, noisy, and spatially sparse measurements [45]. Parallel developments in computational statistics have given rise to hybrid methodologies that integrate stochastic modeling with machine learning [46]. In particular, Neural SDE frameworks [47], physics-informed neural operators (PINO) [48], and operator-learning architectures can now recover the functional structure of the drift f and diffusion G in (1) directly from data, while rigorously enforcing physical and statistical constraints [49]. This fusion of stochastic analysis, numerical approximation, and data-driven inference represents a decisive step toward predictive modeling of complex, high-dimensional, and uncertainty-dominated systems [50].

Over the past decade, several research frontiers have expanded both the theoretical foundations and the computational reach of stochastic modeling [51, 52]. One prominent direction addresses the treatment of high-dimensional and multiscale systems [53], where the number of interacting stochastic degrees of freedom can be in the hundreds or thousands [54], as encountered in turbulent fluid dynamics [55], molecular simulations [56], and eddy-resolving climate models [57]. In such regimes, direct time integration of the full stochastic system is often computationally prohibitive [58]. To overcome this barrier, dimensionality-reduction techniques (such as dynamic mode decomposition, stochastic principal component analysis, and manifold learning [59]) combined with reduced-order and homogenized models [60], have enabled the derivation of effective lower-dimensional stochastic dynamics that preserve essential statistical, spectral, and dynamical properties of the original system while significantly reducing computational cost [61].

Another active research area concerns non-Lipschitz and stiff stochastic dynamics [62], which arise in systems with super-linear drifts, polynomial nonlinearities, or singular coefficients [63]. Classical explicit discretizations, including the Euler-Maruyama scheme [64], may fail in such settings due to numerical instability or divergence. To address these challenges, tamed and balanced numerical integrators have been developed to control the growth of nonlinear terms without sacrificing strong or weak convergence properties [65]. In addition, implicit-explicit (IMEX) schemes [66], exponential integrators [67], and stochastic Rosenbrock-type methods [68] have been successfully adapted to handle severe stiffness and disparate temporal scales, ensuring stability while maintaining high-order accuracy [69].

The modeling of heavy-tailed fluctuations and discontinuous trajectories has likewise progressed significantly [70]. Notable developments include SDEs driven by α-stable Lévy processes [71],

which generate infinite-variance increments and capture heavy-tailed statistics [72]; compound Poisson processes [73], which model finite-activity but potentially large-amplitude jumps [74]; and state-dependent jump intensities [75], allowing the jump rate to vary adaptively with the evolving system state [76]. Such models have become indispensable in quantitative finance [77], where they describe market shocks and volatility clustering; in reliability engineering [78], where they model catastrophic system failures; and in epidemiology, where they capture abrupt changes in transmission due to superspreading events or public-health interventions [79].

Memory-dependent stochastic processes constitute another rapidly advancing frontier [80]. Fractional SDEs and SPDEs [81], incorporating Caputo or Riemann-Liouville derivatives [82], accurately model anomalous diffusion [83], viscoelastic response [84], and long-range temporal correlations observed in diverse physical, biological [85], and socioeconomic systems [86]. The inherently nonlocal structure of fractional operators poses severe computational challenges [87], which have been mitigated by fast convolution algorithms [88], kernel compression strategies [89], and spectral approximations [90], thereby reducing memory and time complexity from $O(N^2)$ to nearly $O(N \log N)$ without significant loss of accuracy [91].

The integration of stochastic modeling with modern machine learning has given rise to a new class of data-driven inference methods [92]. Neural SDEs parameterize the drift and diffusion terms using deep neural networks [93], enabling direct estimation from raw, noisy, or irregularly sampled data while preserving stochastic well-posedness [94]. Operator-learning paradigms [95], including physics-informed neural operators, generalize this framework to SPDEs, learning mappings between infinite-dimensional function spaces under physical and statistical constraints [96]. Furthermore, diffusion-based generative models reinterpret high-dimensional data synthesis as the numerical solution of forward-reverse SDEs, providing a probabilistic bridge between stochastic analysis and state-of-the-art generative learning [97, 98].

Advances in statistical inference have paralleled these methodological innovations [99]. Scalable Bayesian inference for SDE and SPDE models now leverages surrogate modeling [100], Gaussian process emulators [101], and multi-fidelity Monte Carlo schemes to accelerate likelihood evaluations [102]. This has enabled full posterior sampling in previously intractable high-dimensional settings, thereby facilitating rigorous uncertainty quantification in complex stochastic systems [103]. When exact likelihood evaluation is impossible, likelihood-free approaches such as approximate Bayesian computation (ABC) [104] and synthetic likelihood estimation have emerged as powerful alternatives, relying on forward simulation combined with carefully chosen summary statistics [105].

The impact of these developments is evident across multiple disciplines. In epidemiology, stochastic compartmental models augmented with demographic noise [106], Lévy-driven transmission rates [107], and adaptive behavioral feedback mechanisms have achieved superior realism in reproducing epidemic variability [108], persistence probabilities, and extinction events [109]. In quantitative finance, regime-switching jump-diffusion models accurately capture sudden volatility spikes and heavy-tailed return distributions [110, 111]. In climate science, SPDEs driven by spatially correlated colored noise provide a principled framework for representing structured uncertainty in coupled ocean-atmosphere systems [112]. In robotics and control, stochastic optimal control problems (often formulated via backward SDEs) enable robust decision-making under uncertainty, supporting safe trajectory planning and real-time risk-aware operations in dynamic environments [113].

This review delivers a synthesis of recent advances in stochastic modeling with a deliberate emphasis on the *coupled* evolution of (i) stochastic theory (SDEs/SPDEs, including jumps and memory effects), (ii) statistical inference and identifiability under partial/noisy observation, and (iii) algorithmic design principles that determine long-time reliability and practical usability in scientific workflows [114]. While the mathematical foundations of SDEs and SPDEs are well established, the 2020-2025 period has seen a marked consolidation and cross-fertilization of three research streams that are often surveyed *in isolation*: (a) data-driven system identification (e.g., neural SDEs and learned drift-diffusion models), (b) operator learning and physics-informed

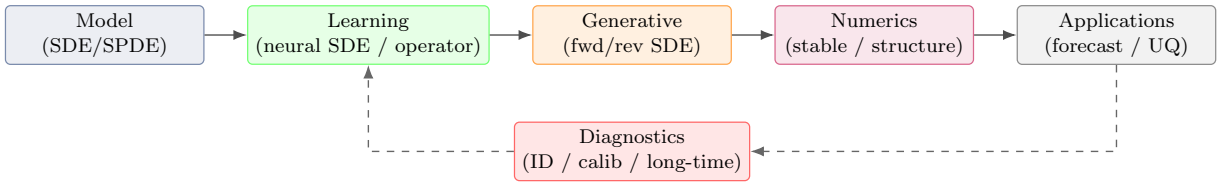


Figure 1. Methodological pipeline.

surrogates for stochastic dynamics (including PINO-type approaches), and (c) generative modeling via forward-reverse SDEs and probability-flow formulations, alongside renewed progress in structure-preserving stochastic numerics. The novel contribution of the present review is therefore not a standalone account of any single strand, but a *unifying taxonomy* that organizes these developments as an end-to-end methodological pipeline from model specification (including heavy tails, jumps, and fractional effects), to learnable representations and inference under incomplete data, to simulation schemes calibrated for stiffness, non-Lipschitz nonlinearities, and fidelity to invariant measures. In addition to summarizing results, we provide critical analysis of where and why methods fail (e.g., drift-diffusion confounding, solver-induced bias, and degradation under sparse/biased biological sampling) and we translate the methodological landscape into application-level guidance by mapping tools to concrete biological and epidemiological tasks such as estimating time-varying transmission or reaction rates, forecasting under reporting artifacts, modeling spatial spread via SPDEs, and quantifying uncertainty for decision support. By consolidating and analyzing works from 2020 to 2025, this review aims to clarify what is now mature, what remains fragile, and which research directions are most promising for building theoretically grounded, computationally scalable, and practically impactful stochastic models in biology and epidemiology [115].

Related surveys typically focus on *one* methodological axis: neural SDE identification and continuous-time latent-variable learning, operator learning/PINOs for PDE-SPDE surrogates, diffusion-modeling and reverse-time SDE sampling, or stochastic numerics for stability and long-time accuracy. These works provide deep coverage within their respective lanes, but they rarely make the *cross-cutting* connections that determine end-to-end performance in scientific applications, for example, how identifiability under partial observation interacts with solver bias, or how generative reverse-time formulations depend on discretization fidelity and data quality. Our review complements these surveys by integrating the four strands into a single pipeline-oriented taxonomy, emphasizing shared failure modes, maturity levels, and minimal evaluation/reporting practices. This perspective is tailored to biomathematics, where practical deployment requires simultaneously credible inference, robust simulation, and interpretable translation to decision-relevant quantities.

As summarized in Fig. 1, our review is organized as a pipeline linking stochastic model specification (SDE/SPDE) to data-driven learning (neural SDEs and operator learning), generative SDE formulations (forward-reverse/score-based diffusion), and structure-preserving numerics for stable simulation and uncertainty quantification in biological and epidemiological applications [93, 96, 97, 161, 162].

The remainder of this article is structured to provide a coherent progression from foundational concepts to advanced applications. Section 2 examines data-driven identification strategies for stochastic systems, with particular attention to neural SDE frameworks and operator-learning methodologies that enable the direct inference of drift and diffusion structures from data. Section 3 explores the emerging role of SDE-based formulations in generative modeling, highlighting their connection to forward-reverse processes and diffusion-based learning. Section 4 offers a detailed survey of numerical integration techniques for SDEs and SPDEs, emphasizing stability, convergence, and long-time accuracy in challenging regimes. Representative applications in biology and epidemiology are presented in Section 5, and the review concludes in Section 6 with a synthesis of open research challenges and promising future directions.

2. Learning stochastic dynamics from data

The growing availability of high-resolution, large-scale datasets has shifted the emphasis in stochastic modeling from prescribing dynamics solely on the basis of first-principles derivations to inferring them directly from empirical observations. Within this data-centric paradigm, two complementary methodologies have emerged. The first addresses finite-dimensional systems through neural SDEs [116], in which both drift and diffusion structures are represented by flexible, trainable mappings estimated from data. The second focuses on infinite-dimensional systems via operator-learning approaches for SPDEs [117], which aim to approximate stochastic evolution operators in a mesh-independent fashion. Both frameworks are designed to assimilate sparse or noisy measurements, generalize robustly across dynamical regimes, and preserve the analytical properties imposed by the underlying stochastic calculus.

2.1. Neural SDEs and drift-diffusion identification

In the neural SDE setting, the drift and diffusion terms are replaced by parameterized functions learned from data, while the governing Itô form is retained [93]. A general formulation on $\mathbb{R}^d$ is

$$\mathrm{d}X_t = f_\theta(X_t, t)\,\mathrm{d}t + G_\theta(X_t, t)\,\mathrm{d}W_t, \qquad X_0 \sim \mu_0, \tag{2}$$

where $X_t \in \mathbb{R}^d$ is the state vector at time t, $f_\theta : \mathbb{R}^d \times [0,T] \to \mathbb{R}^d$ is the drift, $G_\theta : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{d\times m}$ is the diffusion coefficient matrix, θ denotes all trainable parameters, W_t is an m-dimensional standard Wiener process, and μ_0 is the distribution of the initial state. When discrete samples $\{x_{t_k}\}_{k=0}^n$ with time steps $\Delta t_k = t_{k+1} - t_k$ are available, one common identification approach is maximum likelihood under a discretized transition model [118]. Using the Euler-Maruyama approximation [119], the conditional distribution is

$$X_{t_{k+1}} \mid X_{t_k} \approx \mathcal{N}\Big(x_{t_k} + f_\theta(x_{t_k}, t_k)\,\Delta t_k,\ \Sigma_\theta(x_{t_k}, t_k)\,\Delta t_k\Big),$$

where $\Sigma_\theta := G_\theta G_\theta^\top$ is the diffusion covariance. Maximization of this Gaussian likelihood yields an estimate of θ [120]. For state-dependent diffusion, higher-order schemes such as Milstein or Itô-Taylor expansions improve accuracy [121]. Positive semidefiniteness of Σ_θ is enforced via Cholesky factorization, and drift growth is regulated through tamed or monotonicity-preserving architectures to ensure well-posedness and strong solution uniqueness [122].

An alternative approach trains in continuous time by differentiating through the SDE solver. Fixing noise seeds $\xi_k \sim \mathcal{N}(0, I_m)$, a single solver step is denoted by

$$X_{t_{k+1}} = \Phi_\theta(X_{t_k}, t_k; \xi_k),$$

where Φ_θ encodes the numerical update rule [123]. Gradients $\nabla_\theta \mathbb{E}[\mathcal{L}(X_{0:n})]$ are obtained via direct backpropagation or through the adjoint SDE, the latter reducing memory from $\mathcal{O}(n)$ to $\mathcal{O}(1)$ in the number of steps [124]. For partially observed systems $Y_k = HX_{t_k} + \varepsilon_k$, where H is the observation matrix and ε_k is measurement noise, the neural SDE can be coupled with a differentiable filter (such as an ensemble Kalman update or amortized particle weighting) and trained by maximizing a filtering log-likelihood or a variational lower bound [125].

When likelihood-based methods are unreliable, for example under irregular sampling or unknown observation noise, moment-based estimation offers a robust alternative [126]. For any smooth function $g : \mathbb{R}^d \to \mathbb{R}$, Itô's formula gives the local martingale:

$$M_t^g = g(X_t) - g(X_0) - \int_0^t \Big(\nabla g^\top f_\theta + \tfrac{1}{2}\,\mathrm{Tr}[\Sigma_\theta \nabla^2 g]\Big)(X_s, s)\,\mathrm{d}s.$$

Enforcing that empirical averages of $M_{t_k}^g$ vanish yields unbiased estimating equations even with noisy or irregularly spaced data [127]. The diffusion can also be estimated from the quadratic variation:

$$\frac{1}{\Delta t_k}(X_{t_{k+1}} - X_{t_k})(X_{t_{k+1}} - X_{t_k})^\top \approx \Sigma_\theta(X_{t_k}, t_k),$$

with bias corrections applied for finite Δt_k [128]. In models with jumps, drift inference may employ Girsanov's theorem [129], while the jump component can be estimated from its Lévy-Khintchine representation [130], using, for example, compound Poisson thinning or variational approximations for infinite-activity cases [131].

Careful treatment of identifiability is essential: without constraints, drift effects can be spuriously absorbed into the diffusion term [132]. Common remedies include imposing structural restrictions on G_θ (e.g., diagonal or low-rank form), penalizing deviations $\|\widehat{\Sigma}-\Sigma_\theta\|$ over short time increments, or leveraging stationary distributions [133]. If π is the invariant measure of (2), f_θ and Σ_θ satisfy the stationary Fokker-Planck equation

$$\nabla\cdot\big(f_\theta\pi\big) - \tfrac{1}{2}\sum_{i,j}\partial_i\partial_j\big(\Sigma_{\theta,ij}\pi\big) = 0,$$

which can serve as a weak constraint when long-run data are available [134]. In practice, training objectives often combine the transition likelihood, quadratic-variation penalties, weak Fokker-Planck constraints, and stability regularizers that enforce linear growth bounds or one-sided Lipschitz conditions to promote ergodicity [135].

Uncertainty quantification may be addressed via Bayesian neural SDEs [136], where priors are placed on θ and posteriors are explored using stochastic-gradient MCMC or ensemble variational inference [137], or by nonparametric bootstrap resampling over both trajectories and noise seeds [138]. Robustness to out-of-distribution states can be enhanced by spectral normalization in the drift and by imposing lower bounds on diffusion coefficients [139]. In controlled systems with exogenous inputs u_t, the drift $f_\theta(x,t,u)$ can be parameterized in a control-affine form, enabling integration with stochastic optimal control through Hamilton-Jacobi-Bellman surrogates or pathwise policy-gradient methods [140].

2.2. Operator learning for SPDEs (PINO, Neural Operators)

Stochastic partial differential equations on spatial domains $\Omega \subset \mathbb{R}^p$ can be expressed in abstract form as

$$\mathrm{d}u(t) = \mathcal{A}_\phi\big(u(t)\big)\,\mathrm{d}t + \mathcal{B}_\phi\big(u(t)\big)\,\mathrm{d}W_t, \qquad u(0)=u_0, \quad t\in[0,T], \tag{3}$$

where $u(t):\Omega\to\mathbb{R}^q$ denotes the state variable at time t, $\mathcal{A}_\phi$ is a (possibly nonlinear) differential operator parameterized by ϕ, $\mathcal{B}_\phi$ is a noise-multiplication operator, and W_t is a cylindrical Wiener process on an appropriate Hilbert space. The initial condition is $u_0\in\mathcal{H}$, where $\mathcal{H}$ is the state space [141]. The learning objective is to approximate the solution operator

$$\mathcal{S}:(u_0,f,\xi)\mapsto u(\cdot),$$

which maps initial data u_0, deterministic forcing terms $f:\Omega\times[0,T]\to\mathbb{R}^q$, and stochastic realizations ξ to full trajectories u [142]. Neural operator architectures approximate $\mathcal{S}$ directly in function space, enabling mesh-independent generalization and evaluation across discretizations [96]. Two principal designs have been particularly effective.

The first class, Fourier or kernel neural operators, alternates local linear mappings with global integral transforms of the form

$$v_{k+1} = \sigma\big(Wv_k + \mathcal{K}_\psi v_k\big),$$

where v_k denotes the feature representation at layer k, W is a local (pointwise) linear transformation, σ is a nonlinear activation, and $\mathcal{K}_\psi$ is a global integral operator with kernel $K_\psi(x,y)$ parameterized either in physical space or via truncated Fourier multipliers [143]. Stacking such layers yields a mapping $\mathcal{G}_\Psi:\mathcal{X}\to\mathcal{Y}$ that approximates $\mathcal{S}$. In the stochastic setting, randomness enters through sampled forcings f or noise realizations ξ, and the training objective minimizes trajectory- or distribution-based losses over ensembles of noise samples [144]. Stability is promoted by constraining spectral multipliers (e.g., bounding spectral radius) and incorporating skip connections designed to mimic the dissipative structure of the underlying PDE [145].

The second class, PINO, augments data-driven learning with weak enforcement of the governing equations in (3) [146]. Given a set of test functions $\{\varphi_j\}_{j=1}^J\subset\mathcal{H}$, the variational residual is

defined by

$$\mathcal{R}_\phi(u) := \sum_{j=1}^{J} \Big\langle \partial_t u - \mathcal{A}_\phi(u),\ \varphi_j \Big\rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the $L^2(\Omega)$ inner product. Stochastic forcing is incorporated either by sampling noise paths ξ and enforcing Itô weak identities, or by matching statistical quantities (e.g., covariances, energy spectra) derived from theory [147]. The composite training objective takes the form

$$\min_{\Psi,\phi}\ \mathbb{E}_\xi\big[\|\mathcal{G}_\Psi(u_0, f, \xi) - u\|_{\mathcal{Y}}^2\big] + \alpha\, \mathbb{E}_\xi\big[\|\mathcal{R}_\phi(\mathcal{G}_\Psi(u_0, f, \xi))\|^2\big] + \beta\, \mathcal{J}_{\mathrm{BC/IC}},$$

where $\|\cdot\|_{\mathcal{Y}}$ denotes the L^2 norm over $\Omega \times [0, T]$, $\mathcal{J}_{\mathrm{BC/IC}}$ enforces boundary and initial conditions, and $\alpha, \beta > 0$ are weighting parameters [148]. This formulation reduces data requirements and mitigates nonphysical extrapolation in regimes where stochastic forcing excites unresolved spatial or temporal scales.

Accurate noise representation is critical. A common approach is to expand $\xi(t, x)$ in a Karhunen-Loève basis $\{\phi_i(x)\}_{i=1}^r$ [149],

$$\xi(t, x) = \sum_{i=1}^{r} a_i(t)\, \phi_i(x),$$

where $a_i(t)$ are temporally stochastic coefficients and the ϕ_i are spatial modes, possibly learned jointly with the operator [150]. For multiplicative noise, maintaining Itô-Stratonovich consistency requires either explicit correction terms in the residual or learned commutator terms that approximate the conversion [151]. If an invariant measure π is known or estimable, an auxiliary ergodic loss

$$\mathbb{E}_\pi[\mathcal{Q}(u)] \approx \frac{1}{T} \int_0^T \mathcal{Q}\big(\mathcal{G}_\Psi(u_0, f, \xi)(t)\big)\, \mathrm{d}t$$

is used to align surrogate long-time statistics with those of the true system, where $\mathcal{Q}$ denotes a chosen diagnostic such as energy, enstrophy, or a spectral functional [152].

Generalization across meshes and domains is supported by spectral parameter tying to ensure resolution scalability, coordinate-free message passing for unstructured meshes, and discrete conservation constraints such as divergence-free enforcement for incompressible flows or positivity for density fields [153]. For stiff or multiscale SPDEs, multi-rate operator architectures couple coarse global updates with local fine-scale correctors, analogous to heterogeneous multiscale methods [154]. Uncertainty quantification can be incorporated via Bayesian priors on operator weights, ensembles over noise realizations [155], or linearization-based covariance approximations using Gauss-Newton methods [156].

A hybrid approach couples finite-dimensional latent SDEs with learned spatial bases, representing the solution as $u(t, x) \approx \sum_{i=1}^r c_i(t)\, \phi_i(x)$, where $c(t) \in \mathbb{R}^r$ evolves under a neural SDE and $\{\phi_i\}$ are trainable spatial modes [157]. This yields interpretable low-rank stochastic dynamics with mesh-independent inference and efficient sampling, applicable to tasks such as stochastic control, data assimilation, and probabilistic forecasting [158].

Table 1. Critical view of learning stochastic dynamics from data (Sec. 2).
Compact summary of advantages and limits to guide method selection and reporting.

Method	What is learned	Key advantages	Key limits / failure modes	When to use
Neural SDEs and drift-diffusion identification				
MLE (EM) [93, 118–120]	Learn (f_θ, G_θ) via Gaussian transition likelihood (Euler-Maruyama).	Simple, scalable; transparent objective; easy PSD/regularization.	Discretization bias for coarse Δt; brittle under irregular sampling/noise; confounding possible.	Dense/moderate sampling; benchmarking, prototyping.

Continued on next page

Method	**What is learned**	**Key advantages**	**Key limits / failure modes**	**When to use**
Higher-order [121]	Milstein / Itô-Taylor for state-dependent diffusion.	Less bias; more efficient; better for multiplicative noise.	More assumptions/complexity; still sensitive to partial observation.	When EM underfits transitions; multiplicative noise.
Pathwise / adjoint [123, 124]	Differentiate through solver (or adjoint) for task losses.	Flexible beyond likelihood; memory savings; integrates control/forecasting.	Noisy gradients; solver-induced bias; stiff/non-Lipschitz instability.	Long-horizon forecasting/control; non-likelihood learning.
Partial obs + filter [125]	Learn latent SDE from $Y_k = HX_{t_k} + \varepsilon_k$ via differentiable filtering.	Realistic for biology; principled link to observation model.	Filter degeneracy/linearization bias; identifiability worsens.	Hidden states; sensor noise; irregular sampling.
Martingale moments [126, 127]	Estimate via Itô-martingale identities / estimating equations.	Robust to irregular grids; interpretable constraints; complements MLE.	Test-function choice matters; can be less efficient.	Irregular sampling; uncertain noise models.
Quadratic variation [128]	Estimate Σ_θ from short-time increments (with corrections).	Direct diffusion signal; useful diagnostic vs confounding.	Sensitive to measurement noise; finite-Δt bias.	High-frequency data; auxiliary loss/diagnostic.
Jumps (Lévy) [129–131]	Infer drift+jumps (Girsanov / Lévy-Khintchine / variational thinning).	Captures bursts/heavy tails; relevant for abrupt interventions.	Hard identifiability; masked by noise/discretization; high cost.	Shock-driven/event data; heavy-tailed regimes.
Identifiability + stationarity [132–135]	Structure on G_θ + weak Fokker-Planck / ergodic constraints.	Targets drift-diffusion confounding; improves scientific credibility.	Stationarity may fail; PDE constraints stiff; risk underfitting transients.	Long time series; quasi-stationary regimes.
UQ (Bayes/ens.) [136–139]	Posterior/ensemble/bootstrap over θ and noise seeds.	Calibrated uncertainty; supports decisions; detects overconfidence.	Expensive; calibration degrades OOD; needs reporting standards.	Forecasting/policy; when uncertainty is critical.
Operator learning for SPDEs (Neural Operators, PINO)				
Neural operators [96, 142–145]	Learn solution operator $\mathcal{S} : (u_0, f, \xi) \mapsto u(\cdot)$.	Mesh-independent surrogates; fast sampling/UQ after training.	OOD failures (regimes/geometries); stability not automatic; nonphysical outputs.	Multi-query emulation; UQ/optimization loops.
PINO [146–148]	Data loss + weak SPDE residual (variational constraint).	Less data; reduces nonphysical extrapolation; residual diagnostics.	Residual weighting delicate; rare-event/tails hard; Itô/Strat subtle.	Known physics + limited labels; extrapolation-sensitive tasks.
Noise basis (KL) [149, 150]	Low-rank noise expansions; learn modes jointly.	Tractable randomness; sample-efficient; interpretable modes.	Rank truncation bias; non-identifiability; tail distortion.	Dominant-mode stochasticity; moderate dimensions.
Itô/Strat fixes [151]	Consistency corrections for multiplicative noise.	Reduces structural bias; improves transfer.	Implementation burden; often neglected (silent error).	Multiplicative-noise SPDEs; sensitive physics.
Ergodic loss [152]	Match invariant-measure diagnostics (energy/moments/spectra).	Improves distributional credibility; supports risk endpoints.	Needs long runs/robust estimators; may trade short-term accuracy.	Long-horizon stats; tail-risk applications.
Mesh/domain gen. [153, 154]	Coordinate-free design + constraints (positivity, divergence-free); multirate.	Portability; fewer artifacts; helps stiff/multiscale regimes.	Engineering effort; constraints reduce flexibility; tuning issues.	Irregular meshes; conservation structure; stiffness.
UQ for operators [155, 156]	Bayes/ensembles or covariance via linearization (GN).	Credible uncertainty propagation; supports design decisions.	OOD calibration hard; Gaussian approx misses tails; cost.	High-stakes prediction; limited SPDE truth.
Hybrid latent SDE [157, 158]	Low-rank $u(t,x) \approx \sum c_i(t)\phi_i(x)$ with neural SDE for $c(t)$.	Interpretable; efficient sampling; mesh-independent inference.	Rank choice critical; misses localized/rare extremes.	Coherent-structure systems; fast probabilistic prediction.

Table 1 highlights a maturity gradient across the learning pipeline. For neural SDEs, likelihood- and solver-based training are now well established, but *identifiability* (drift-diffusion confounding)

and *solver-induced bias* remain the main blockers for reliable scientific inference under partial observation [124, 132, 135]. For operator learning, mesh-independent surrogates have matured rapidly, yet the dominant open challenges are *out-of-distribution generalization* and *long-time statistical fidelity*, especially under multiplicative noise and rare-event regimes [145, 151, 152]. These observations motivate the reporting standards proposed later (distributional scores, coverage diagnostics, and invariant-measure checks) as minimal safeguards for biological deployment.

3. Generative Modeling as SDEs

Generative modeling aims to construct stochastic transformations that map a tractable reference distribution onto a complex target distribution [24]. Recent advances have established that score-based generative models and denoising diffusion models can be formulated entirely within the stochastic differential equation (SDE) framework introduced in (1) [97]. In this setting, a *forward* process progressively perturbs samples from the data distribution p_{data} into a simple reference law p_{ref}, typically a standard Gaussian, over a fixed time horizon [159]. A *reverse-time* stochastic process then reconstructs new samples by inverting this diffusion [160], as illustrated in Fig. 2.

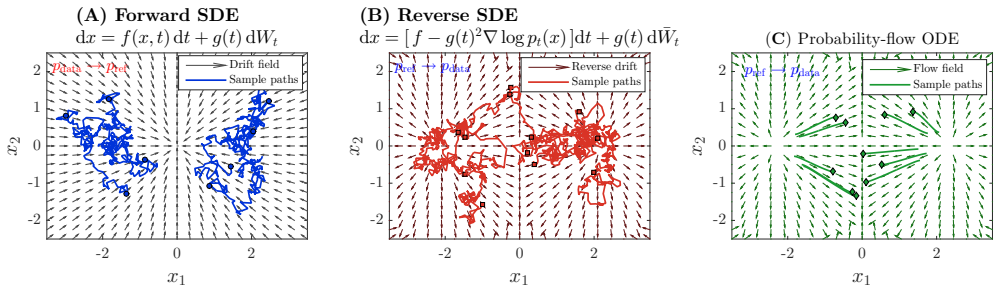


Figure 2. Forward-reverse SDE framework and its deterministic probability-flow ODE counterpart. Panel (A) shows the forward SDE transporting p_{data} to a reference Gaussian p_{ref}. Panel (B) shows the reverse SDE mapping p_{ref} back to p_{data} via the learned score function $\nabla \log p_t(x)$. Panel (C) depicts the equivalent probability-flow ODE sharing the same marginals as the forward and reverse SDEs but evolving deterministically. Vector fields represent drift components; colored trajectories indicate sample paths.

Let p_t denote the marginal distribution of the forward process at time t. Under regularity conditions, Anderson's time-reversal theorem [161] yields the reverse-time SDE:

$$\mathrm{d}x(t) = \Big[f(x(t),t) - g(t)^2 \, \nabla_x \log p_t\big(x(t)\big) \Big] \, \mathrm{d}t + g(t) \, \mathrm{d}\bar{w}(t), \tag{4}$$

where $\bar{w}(t)$ is a standard Wiener process evolving backward in time, f and g are the drift and diffusion coefficients of the forward dynamics, and $\nabla_x \log p_t$ is the score function of p_t. Since p_t is not known in closed form, the score function is approximated by a trainable surrogate $\hat{s}_\varphi(x,t)$, estimated through score-matching or denoising-score-matching objectives, ensuring asymptotic consistency in the limit of infinite data and model capacity [162].

Numerical generation is performed by discretizing (4) with schemes such as Euler-Maruyama [163], stochastic Runge-Kutta [164], or predictor-corrector integrators [165]. The choice of scheme directly influences sampling bias and variance, with high-order solvers yielding greater fidelity for a given computational budget [166]. An important alternative is the *probability-flow ordinary differential equation* associated with the forward process, which evolves deterministically yet preserves the same marginal distributions as the reverse SDE, often permitting larger stable time steps [167].

In scientific applications, the forward dynamics can be tailored to preserve statistical invariants, enforce spectral constraints, or emulate dissipative mechanisms from coarse-grained physics [168]. The learned reverse process then reconstructs fine-scale structure and corrects higher-order moments, enabling physically consistent synthesis in high-dimensional contexts such as turbulent

flows, molecular dynamics, and climate modeling [169]. Computational efficiency can be further enhanced through adaptive time stepping, variance-reduction strategies (e.g., control variates), and reduced-order latent representations [170].

A natural generalization introduces exogenous control inputs u_t into the drift, yielding a controlled reverse-time process of the form

$$f(x, t, u_t) - g(t)^2 \nabla_x \log p_t(x, t),$$

which connects generative modeling with stochastic optimal control [171]. In this interpretation, the reverse dynamics act as an optimal feedback policy minimizing a cost functional equivalent to the negative log-likelihood, enabling targeted sampling subject to trajectory or terminal constraints [172]. This formulation has direct implications for data assimilation, scenario generation, and constraint-aware simulation [173].

By embedding generative modeling entirely within the SDE formalism and incorporating domain-specific constraints, one obtains a unifying mathematical structure capable of preserving statistical integrity while extending these methods from static data generation to the synthesis and control of complex stochastic dynamical systems.

Table 2. Critical view of generative modeling as SDEs (Sec. 3). Compact summary of the forward-reverse SDE pipeline, the main solver choices, and the dominant failure modes (score error compounding, discretization bias, tail miscalibration, and constraint-induced bias) relevant for scientific applications.

Component / method	Core idea / what is learned	Key advantages (critical view)	Key limits / failure modes	When to use
Forward diffusion and reverse-time sampling				
Forward SDE (data → ref) [24, 97, 159]	Defines a noising process transporting p_{data} to a tractable p_{ref} (often Gaussian) over $t \in [0, T]$.	Mathematically clean bridge between diffusion models and SDEs; controllable noise schedule; enables likelihood/ELBO-style training.	Poor schedule choice induces ill-conditioned learning (vanishing/exploding SNR); may destroy relevant invariants; can amplify dataset bias.	Generic high-dimensional generation; as a baseline pipeline before adding constraints.
Reverse-time SDE [160, 161]	Samples by integrating the time-reversed SDE with drift correction via the score $\nabla_x \log p_t(x)$.	Flexible sampling; supports conditional generation; can incorporate domain constraints via drift modifications.	Sensitive to score error (compounds backward); discretization bias; numerical instability for stiff dynamics.	When exact sampling is not required but high fidelity is; conditional/scenario generation.
Score estimation (DSM / SM) [162]	Learns $\hat{s}_\varphi(x, t) \approx \nabla_x \log p_t(x)$ via score matching or denoising score matching.	Avoids explicit density evaluation; scalable to high dimension; asymptotically consistent (idealized limit).	Finite-data/model bias; poorly calibrated tails/rare events; sensitive to measurement noise and covariate shift.	Large datasets; high-dimensional fields; when density evaluation is infeasible.
Numerical solvers and sampling accelerations				
Euler-Maruyama [163]	First-order discretization of reverse SDE; simplest sampler.	Fast, simple; robust baseline; easy to implement for benchmarking.	Large step-size bias; many steps needed for quality; unstable if drift/score is stiff.	Quick experiments; low-budget runs; baseline comparisons.
Higher-order SDE solvers (SRK) [164, 166]	Stochastic Runge-Kutta / higher-order discretizations to reduce bias/variance per step.	Better fidelity per compute; fewer steps for same quality; improved stability in moderate stiffness.	More complex; requires additional evaluations; gains depend on regularity assumptions.	When EM is too biased/slow; medium-to-high accuracy sampling.

Continued on next page

Component / method	Core idea / what is learned	Key advantages (critical view)	Key limits / failure modes	When to use
Predictor-corrector [165, 166]	Alternates prediction with score-based correction (e.g., Langevin-type refinements).	Often improves sample quality at fixed steps; can reduce discretization artifacts; practical robustness.	Heuristic tuning; can increase compute; may over-smooth or mishandle sharp features.	Image/field synthesis; practical high-quality sampling under limited steps.
Probability-flow ODE [167]	Deterministic ODE sharing the same marginals as the SDE; enables ODE solvers/adaptive stepping.	Deterministic trajectories; often larger stable steps; easier likelihood computation in some settings.	ODE discretization error still matters; can lose stochastic diversity if not handled carefully; may be sensitive to solver tolerances.	Fast sampling with adaptive solvers; when determinism and reproducibility matter.
Adaptive stepping + variance reduction [170]	Controls local error and reduces variance via control variates / adaptive step schedules / reduced-order latents.	Efficiency gains for stiff or multiscale systems; better compute-quality tradeoff.	Complexity and tuning; may break guarantees if heuristics dominate; risk of biased estimates.	Scientific simulation surrogates; multiscale/stiff dynamics; compute-constrained workflows.
Scientific constraints, control, and reliability				
Physics-consistent reconstruction [168, 169]	Reverse process restores fine scales and higher-order statistics under constraints (moments, spectra, conservation).	Better long-time/statistical fidelity; improves interpretability for scientific use.	Hard for rare events and tails; constraint enforcement may be approximate; evaluation must be distributional.	Scientific field synthesis; uncertainty propagation; scenario generation with plausibility constraints.
Controlled generative SDEs / control link [171–173]	Introduces control inputs u_t in the drift; interprets reverse dynamics as feedback policy for targeted sampling.	Connects diffusion models to stochastic control and data assimilation; enables constrained/goal-conditioned generation.	Control objectives can conflict with fidelity; sensitive to model bias; stability under constraints can be nontrivial.	Data assimilation; constraint-aware scenarios; goal-directed sampling under terminal/path constraints.

Table 2 indicates that the dominant bottleneck is typically *not* the forward SDE itself but the interaction between (i) score-estimation error (especially in tails/rare events), and (ii) discretization error that compounds during reverse-time integration [162, 166]. In sparse or biased biological datasets, these effects can manifest as miscalibrated uncertainty and overconfident extrapolation, motivating distributional diagnostics (e.g., coverage, CRPS/PIT) and sensitivity checks to solver tolerances as minimal safeguards for deployment [167, 170].

4. Numerical methods for SDEs and SPDEs

The accurate and efficient numerical integration of stochastic differential equations and their infinite-dimensional counterparts is central to both predictive modeling and statistical inference. The numerical discretization must preserve the fundamental properties of the underlying stochastic dynamics, including convergence order, stability, invariant measures, and qualitative behavior over long time horizons. This section surveys recent methodological developments across four interconnected areas.

4.1. Strong and Weak Convergence, and Stability

In the numerical treatment of stochastic differential equations, strong convergence quantifies the mean-square error between the numerical approximation $X_T^{\Delta t}$ and the exact solution X_T at a fixed terminal time T [174],

$$\big(\mathbb{E}\,\|X_T - X_T^{\Delta t}\|^2\big)^{1/2} = \mathcal{O}\big((\Delta t)^{\gamma}\big),$$

while weak convergence measures the error in expectations of sufficiently smooth observables φ,

$$\left|\mathbb{E}\,\varphi(X_T) - \mathbb{E}\,\varphi(X_T^{\Delta t})\right| = \mathcal{O}\big((\Delta t)^{\beta}\big).$$

For drift-diffusion systems with non-globally Lipschitz drift, polynomial growth, or multiplicative noise, modern analyses derive orders γ and β under one-sided Lipschitz and coercivity conditions, often with tamed, balanced, or implicit updates to ensure well-posedness of the discrete dynamics [62,175,176]. In infinite dimensions, the same notions apply to semidiscrete SPDEs in Hilbert spaces, where temporal error interacts with spatial discretization error through the smoothing properties of the linear generator and the regularity of the noise [177]. Stability considerations extend beyond mean-square boundedness to include almost-sure exponential stability [178], boundedness in probability, and preservation of qualitative invariants over long horizons [179]. Empirical slopes for representative strong and weak errors, and their asymptotic rates, are illustrated in Fig. 3.

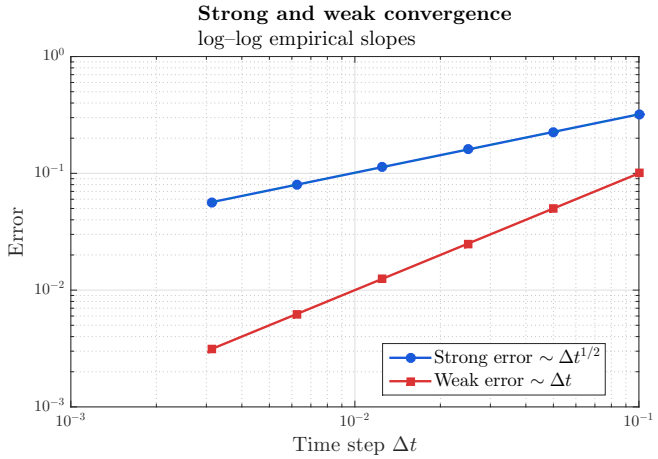


Figure 3. Strong and weak convergence on a log-log scale for representative schemes. The strong error exhibits the expected order one half for an Euler-Maruyama discretization with multiplicative noise, whereas the weak error attains order one for smooth test functionals. Dashed guides indicate the asymptotic slopes used to assess rates. The same methodology extends to semidiscrete SPDEs, where temporal and spatial errors interact through the smoothing properties of the linear part and the regularity of the stochastic forcing.

4.2. Tamed, implicit, and balanced schemes

Classical explicit schemes may diverge when applied to SDEs with superlinear drift growth. Tamed schemes modify the drift term to control growth without sacrificing explicitness [180]. For instance, a tamed Euler method replaces $f(x)$ by $f(x)/(1 + \alpha\Delta t\|f(x)\|)$, ensuring boundedness of increments while preserving strong convergence of order one under appropriate conditions [181]. Implicit schemes, such as backward Euler-Maruyama, provide improved stability properties, particularly for stiff stochastic systems and SPDEs with dissipative operators [182]. Balanced methods interpolate between explicit and implicit updates, adjusting drift and diffusion contributions to simultaneously control stability and accuracy [183]. These approaches have been extended to higher-order stochastic Runge-Kutta methods and to adaptive time-stepping algorithms that refine resolution in regions of high stochastic activity [184].

4.3. Long-time accuracy and invariant measures

For ergodic SDEs, a central requirement on a numerical method is the faithful approximation of the invariant law over long horizons [185]. If π denotes the stationary distribution of the exact process, a method with step size Δt induces a Markov chain with stationary law $\pi^{\Delta t}$ [186]. Long-time accuracy entails weak convergence $\pi^{\Delta t} \Rightarrow \pi$ as $\Delta t \to 0$ and a bias in ergodic averages that

decays at an optimal rate [187]. The asymptotic bias of an observable $\mathcal{Q}$ admits Poisson-equation representations that expose its dependence on the generator and the local truncation error [188]; this viewpoint underpins modified-equation corrections and Talay-Tubaro-type expansions for invariant-measure error [189]. In practice, drift and diffusion adjustments, symmetrization, and schemes that preserve reversibility or detailed balance reduce long-time bias while retaining acceptable finite-time accuracy [190]. In infinite dimensions, structure-preserving time integrators for SPDEs target invariants and statistical diagnostics dictated by the physics, such as energy spectra [191], enstrophy [192], or mass [193], and are analyzed using spectral gaps or Wasserstein-contractivity of the associated Markov semigroups [194]. Figure 4 illustrates the approach of time averages toward the invariant expectation and highlights the reduced residual bias of an implicit method relative to explicit Euler-Maruyama at a fixed step size.

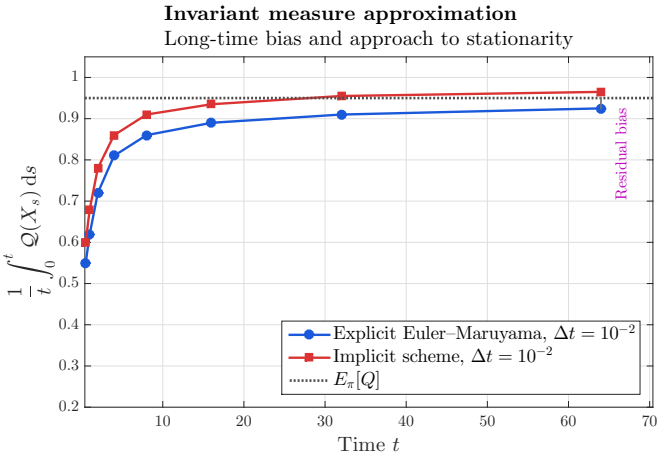


Figure 4. Convergence of time averages to the invariant expectation for an ergodic observable $\mathcal{Q}$. Curves show $\frac{1}{t}\int_0^t \mathcal{Q}(X_s)\,\mathrm{d}s$ for an explicit Euler-Maruyama discretization and an implicit method at the same step size. The dotted line marks $\mathbb{E}_\pi[\mathcal{Q}]$, and the vertical segment at the final time indicates the residual bias. The implicit scheme approaches the stationary value more rapidly and with smaller long-time bias, consistent with theoretical predictions based on generator perturbations and Poisson-equation error analysis.

4.4. Multilevel Monte Carlo and Rare-Event simulation

Multilevel Monte Carlo reduces work by coupling approximations across a hierarchy in time and, when relevant, space [195]. Let φ_ℓ denote an approximation of $\varphi(X_T)$ at level ℓ with step Δt_ℓ (and mesh width h_ℓ if spatial discretization is present). The telescoping identity:

$$\mathbb{E}[\varphi_L] = \mathbb{E}[\varphi_0] + \sum_{\ell=1}^{L} \mathbb{E}\big[\varphi_\ell - \varphi_{\ell-1}\big]$$

is estimated by Monte Carlo averages with N_ℓ *coupled* samples [196]. Three rates govern complexity: the weak bias $\mathcal{O}(\Delta t_L^{\alpha})$, the variance decay $\mathrm{Var}(\varphi_\ell - \varphi_{\ell-1}) = \mathcal{O}(\Delta t_\ell^{\beta})$, and the per-sample cost $\mathcal{O}(\Delta t_\ell^{-\gamma})$. With the optimal allocation $N_\ell \propto \sqrt{\mathrm{Var}(\varphi_\ell - \varphi_{\ell-1})/C_\ell}$,

$$\mathcal{C}_{\mathrm{MLMC}} = \begin{cases} \mathcal{O}(\varepsilon^{-2}), & \beta > \gamma, \\ \mathcal{O}(\varepsilon^{-2}\log^2 \varepsilon), & \beta = \gamma, \\ \mathcal{O}\big(\varepsilon^{-2-(\gamma-\beta)/\alpha}\big), & \beta < \gamma, \end{cases}$$

which improves on single-level sampling whenever $\beta \geq \gamma$. For Lipschitz SDEs with Euler-Maruyama couplings one typically has $\alpha = 1$, $\beta \approx 1$, and $\gamma \approx 1$, leading to $\mathcal{O}(\varepsilon^{-2}\log^2 \varepsilon)$ [197]. Antithetic constructions or Milstein-type ideas without Lévy areas often restore $\beta > \gamma$ and the canonical $\mathcal{O}(\varepsilon^{-2})$ [198]. In SPDEs, multilevel couplings combine temporal integrators with

spectral or finite-element hierarchies; multi-index Monte Carlo extends the telescoping structure across time-space refinement directions [199].

Rare-event probabilities demand variance control beyond stratification by level [200]. Importance sampling with optimally tilted measures derived from large-deviation asymptotics, splitting or subset simulation with adaptively placed intermediate thresholds, and adaptive multilevel splitting or interacting particle systems yield estimators that remain efficient as the probability decreases [201]. These mechanisms integrate naturally with the multilevel decomposition by biasing each difference $\varphi_\ell - \varphi_{\ell-1}$, by coupling biased path pairs, or by embedding control-based surrogates of the optimal change of measure [202]. The pipeline in Fig. 5 summarizes this workflow from hierarchy construction, through telescoping and estimator assembly, to variance-reduction layers for the tail regime.

Overall performance reflects the balance between bias order, coupling variance, long-time statistical fidelity, and variance-reduction design. In practice, structure-preserving or higher-order discretizations are paired with MLMC (or multi-index MLMC) for bulk uncertainty quantification, while importance sampling, splitting, or adaptive multilevel splitting is activated selectively for tail-sensitive observables.

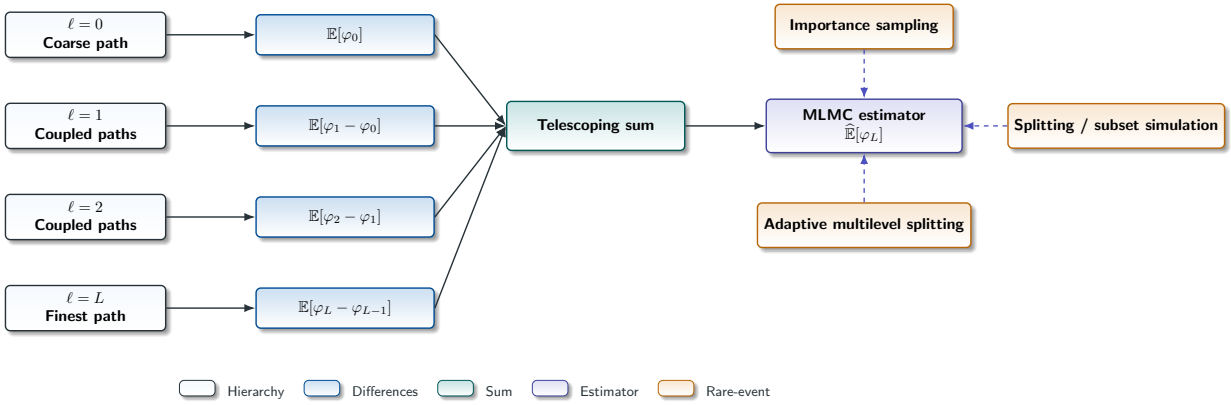


Figure 5. Schematic representation of the MLMC pipeline augmented with rare-event variance-reduction strategies.

5. Applications in biological and epidemiological modeling

To strengthen practical relevance, we synthesize the pipeline from Sections 2-3 into a reproducible workflow for biological and epidemiological studies, where data are typically sparse, biased, and partially observed [203–209]. Model choice is scale-driven: finite-dimensional compartmental / reaction-network systems support neural-SDE identification for time-varying rates, whereas spatial spread and dispersal motivate SPDE formulations and operator-learning surrogates (Section 2.2). Because under-reporting, reporting delays, and heterogeneous testing form an explicit noisy measurement channel, latent-state learning should be coupled with differentiable filtering/assimilation (Section 2.1) rather than trained directly on raw case counts. Reliable inference then combines transition-likelihood or pathwise objectives with quadratic-variation and martingale moment constraints, supplemented (when long time series permit) by weak stationarity or invariant-measure checks to mitigate drift-diffusion confounding and solver-induced bias. Forecasting and uncertainty quantification follow via ensemble simulation, Bayesian neural SDEs, or calibrated bootstrap procedures, evaluated using both pointwise and distributional criteria (e.g., CI coverage and distributional scores) to avoid overconfident extrapolation. Finally, scenario generation and counterfactual analysis can be posed within the generative SDE framework (Section 3), where controlled reverse-time dynamics enable constraint-aware synthesis while preserving target marginals, and operator-learning surrogates provide computationally efficient sampling for spatial forecasts.

Table 3 maps the method families in Sections 2-4 to biological/epidemiological use-cases (2020-2025): (A) neural SDE identification, (B) operator learning and physics-informed PDE/SPDE surrogates, (C) reverse-time generative SDEs, (D) structure-preserving stochastic numerics, and (E) inference/surrogates/phenomena. An entry ✓ denotes at least one relevant published work or public preprint; ✗ denotes a gap (no result found as of August 2025). Across domains, the central objectives are consistent: estimation of time-varying transmission/reaction rates, partial-observation inference, preservation of positivity/invariants, and uncertainty propagation for policy-oriented forecasting.

On the SDE side, neural SDEs provide flexible yet mechanistically grounded parameterizations for noisy biological dynamics and epidemic time series [210], with discretized likelihood calibration via Euler-Maruyama remaining a practical default for discretely sampled outbreaks [221]. When stiffness and signal-to-noise demand higher fidelity, Milstein/Itô-Taylor/SRK schemes improve simulation-and-fit efficiency [217]. Positive-semidefinite diffusion parameterizations (e.g., Cholesky factors) stabilize multivariate learning [222], while tamed or monotonicity-preserving architectures mitigate superlinear drift pathologies arising from saturating incidence or autocatalytic kinetics [223]. Continuous-time training through differentiable SDE solvers has been explored in biophysical and biopharmaceutical identification tasks [224]. In practice, partial and biased observation is the rule, motivating differentiable filters (EnKF/particle) for joint state-parameter learning [225] and martingale/moment estimators as complementary, interpretable checks [226]. Bayesian neural SDEs (e.g., SG-MCMC and variational ensembles) deliver posterior trajectories and calibrated intervals for epidemic indicators [227].

When spatial coupling, diffusion, and nonlocal effects dominate, operator learning extends the toolkit beyond low-dimensional SDEs. Neural operators (Fourier/kernel variants) and PINN/PINO-style training have been adapted to SPDE reaction-diffusion and metapopulation settings, improving extrapolation while retaining mechanistic constraints [228, 229]. Long-time alignment through ergodic/statistical losses helps match stationary behavior in biochemical and ecological networks [230], multirate/multiscale architectures address separated time scales [231], and hybrid low-rank constructions (learned spatial bases with latent temporal SDEs) yield compact surrogates for spatially aggregated epidemics [232]. By contrast, rigorous mesh/domain generalization with explicit conservation/positivity guarantees and operator-level Bayesian UQ remains underreported (✗).

Generative SDEs provide complementary capability: score-based diffusion models enable strong priors and data augmentation for biological trajectories and epidemic curves [233], and predictor-corrector samplers support accurate path synthesis and parameter exploration in low-data regimes [234]. Open opportunities remain in controlled reverse-time formulations linked to stochastic optimal control, variance reduction, and reduced-order sampling under mechanistic constraints (✗).

Credible forecasting depends on discretizations that respect biological structure-positivity of compartments, invariance of feasible sets, and stability under stiffness. Recent studies establish strong/weak convergence and stability for SDE/SPDE schemes used in epidemiology, including linearly implicit and explicit positivity-preserving integrators [216, 218, 220]. For stochastic reaction-diffusion SPDEs, unconditionally stable backward-Euler/implicit finite-difference schemes support practical step-size rules [212], and explicit positivity-preserving methods can provide extinction guarantees in subcritical regimes [220]. Still underexplored in applied pipelines are tamed/balanced schemes for superlinear drifts, adaptive time stepping, and multilevel Monte Carlo for forward UQ (✗), despite clear relevance to bursty transmission and regime shifts.

Finally, modern inference and surrogates complete the application pipeline: sequential Monte Carlo supports online calibration under delays and regime changes [215]; likelihood-free neural posterior estimation amortizes inference across mechanistic stochastic epidemic simulators for rapid scenario analysis [211]; neural surrogates for stochastic SIS-type models reduce computational cost in design/control loops [214]; interacting jump-diffusions and common-noise effects clarify synchronization and variability in metapopulations and environmentally forced systems [213]; and stochastic Runge-Kutta methods remain practical tools for vector-borne dynamics

and multiscale uncertainty propagation [217]. Across single-cell dynamics, pharmacological response, vector-borne and waterborne outbreaks, and respiratory epidemics, four recurring themes emerge: mechanistically informed learning, partial-observation inference, structure-preserving numerics, and scalable UQ/surrogates for decision support-with the most stable pipelines combining identifiable drift parameterizations, PSD-constrained diffusions, filter-based training, and positivity-preserving or implicit time integration.

Table 3: Recent works (2020-2025) applying the surveyed methods to biology and epidemiology. ✓ = at least one study found; ✗ = no results found to date.

Method	Area	Status	Refs	Notes
A. Neural SDE identification and learning				
Neural SDEs (learned drift/diffusion)	Neural SDE	✓	[210]	scDiffEq from single-cell time series.
MLE (Euler-Maruyama transitions)	SDE ID	✓	[221]	Cholera data; discretized likelihood.
Higher-order integrators (Milstein/Itô-Taylor/SRK)	SDE numerics	✓	[217]	Dengue SDE simulations (SRK).
PSD diffusion (Cholesky)	SDE ID	✓	[222]	PSD covariance parameterization.
Tamed / monotone architectures	Robust learning	✓	[223]	Stable learning for nonlinear epidemic dynamics.
Continuous-time training via SDE solvers	System ID	✓	[224]	Pharmacological dynamics (CT-ESN).
Partial obs. + differentiable filters	State-param. inf.	✓	[225]	Particle/EnKF-style inference (monkeypox).
Moment / martingale estimation	Stochastic proc.	✓	[226]	CTMC moments / MFPT bounds.
Bayesian neural SDEs	Bayesian inf.	✓	[227]	COVID-19 posterior dynamics (Germany).
Bootstrap UQ (seeds/paths)	UQ	✗	-	No results found.
B. Operator learning and PINNs for SPDE / PDE models				
Neural operators (Fourier/kernel)	Operator (SPDE)	✓	[228]	Physics-informed fractional epidemic models.
PINO / PINNs	Operator learn.	✓	[229]	PINN-style training for epidemic dynamics.
Ergodic/statistical losses	Long-time align.	✓	[230]	Ergodicity-guided inference.
Mesh/domain gen. + conservation	Operator (SPDE)	✗	-	No results found.
Multirate / multiscale operators	Operator learn.	✓	[231]	Multilevel modeling for comp. epidemiology.
Operator-level UQ	Operator (SPDE)	✗	-	No results found.
Hybrid latent SDE + spatial bases	Hybrid	✓	[232]	Low-rank hybrid for COVID-19.
C. Generative SDEs and reverse-time modeling				
Reverse-time SDE / prob.-flow ODE	Generative SDE	✓	[233]	Score-based reverse-time view.

Continued on next page

Table 3: Recent works (2020-2025) applying the surveyed methods to biology and epidemiology. ✓ = at least one study found; ✗ = no results found to date. (Continued)

Method	Area	Status	Refs	Notes
Predictor-corrector samplers	Generative SDE	✓	[234]	Parameter estimation strategies.
Adaptive/VR/reduced-order sampling	Generative SDE	✗	-	No results found.
Controlled reverse-time (control link)	Generative SDE	✗	-	No results found.
D. Numerical analysis for SDEs / SPDEs				
Strong/weak conv. + stability	Numerics	✓	[216, 218, 220]	Baseline + SIS implicit + positivity schemes.
Implicit SPDE (BE + implicit FD)	SPDE num.	✓	[212]	Unconditionally stable reaction-diffusion.
Linearly implicit SIS schemes	SDE num.	✓	[218]	Stability/accuracy for stiff incidence.
Tamed schemes (superlinear drift)	SDE num.	✗	-	No results found.
Balanced schemes	SDE num.	✗	-	No results found.
Adaptive time stepping	SDE/SPDE num.	✗	-	No results found.
Invariant-measure fidelity	Long-time num.	✓	[216]	Long-horizon bias / invariant measures.
Positivity / invariance preserving	Structure	✓	[220]	Positivity + extinction guarantees.
MLMC / multi-index MLMC	UQ	✗	-	No results found.
E. Inference, surrogates, and phenomena				
Sequential Monte Carlo	Inference	✓	[215]	Online state + time-varying parameter tracking.
Likelihood-free NPE	Inference	✓	[211]	Amortized posteriors for epidemic simulators.
ANN surrogates for stochastic SIS	Surrogates	✓	[214]	LM-trained emulator for SIS.
Jumps / common-noise phenomena	Phenomena	✓	[213]	Jump-diffusion + common-noise variability.
Stochastic Runge-Kutta	SDE num.	✓	[217]	Dengue uncertainty propagation.

6. Outlook and open problems

The convergence of data-driven identification, continuous-time generative modeling, and structure-preserving numerics is reshaping stochastic modeling for biological and epidemiological systems. Table 4 consolidates the most actionable research directions by pairing each open problem with a precise objective, plausible methodological routes, and reportable evaluation criteria.

For neural SDEs under partial observation (especially with state-dependent diffusion) priority must be given to verifiable identifiability and finite-sample guarantees that prevent drift-diffusion confounding and remain robust to irregular sampling. For operator learning, the central requirement is to extend mesh-independent approximation to stochastic forcing while providing stability and well-posedness guarantees (strong solutions, pathwise stability, geometric ergodicity, and, in Bayesian settings, posterior contraction). These guarantees are essential to support principled recovery of dynamics and uncertainty rather than ad hoc curve-fitting (Table 4, "Identifiability" and "Consistency & well-posedness").

Training should also be *discretization-aware*: objectives ought to be optimized under the same integrator used at inference to eliminate train-test discretization mismatch. This motivates

memory-stable adjoints and AD-compatible schemes with strong stability/structure properties (e.g., tamed or balanced Milstein variants, linearly implicit IMEX updates, and positivity-preserving projections), together with in-loop adaptivity that separates discretization error from statistical error during learning (Table 4, "Solver-aware training" and "Adaptive error control").

Finally, evaluation should move beyond pointwise intervals toward calibrated, shift-robust sequential forecasting: distinguish aleatoric from epistemic uncertainty, incorporate dependence-aware conformal calibration, and report distributional diagnostics (CRPS/energy scores, PIT) alongside long-horizon bias in invariant statistics. Rare-event and tail-risk quantification remains comparatively underdeveloped; coupling rare-event biasing with multi-index MLMC is a promising route to predictable variance-cost behavior for exceedance probabilities (Table 4, "Tail risk & rare events" and "Benchmarking & reporting").

Table 4. Open problems distilled into goals, candidate methods, and reportable metrics for learned stochastic dynamics and structure-aware numerics.

Theme	Goal	Candidate methods	Metrics
Identifiability (neural SDE)	Disentangle drift vs. diffusion under partial observation and irregular sampling.	Structured diffusion (factorized/low-rank), short-increment penalties; stationary-moment constraints; uniform empirical-process bounds; confounding tests.	Drift/diffusion MSE; misattribution rate; CI coverage.
Consistency & well-posedness	Guarantee strong solutions, stability, ergodicity, and (Bayesian) posterior contraction.	One-sided Lipschitz & linear-growth priors; Lyapunov drift/minorization; Wasserstein contractivity; sieve priors and small-ball conditions.	Stability constants; mixing rate; ergodic bias; contraction rate.
Solver-aware training	Remove train-inference discretization mismatch.	Discretization-aligned losses; reversible/checkpointed adjoints; step-size-aware objectives; gradient-stability control for stiffness.	Generalization gap (train vs. inference); gradient growth; solver-induced bias.
Structure-preserving SPDE integrators	Maintain invariants (mass/positivity) and spectra over long horizons.	Linearly implicit IMEX; positivity-preserving projections; energy/enstrophy-aware updates; invariant-region enforcement.	Invariant error; spectral discrepancy; long-time bias.
Adaptive error control (in-loop)	Separate discretization vs. statistical error and adapt time/mesh during training.	Dual-weighted residuals; MLMC/MIMC gradient estimators; online mesh/time refinement.	Work-RMSE tradeoff; estimator variance; speedup at fixed error.
Tail risk & rare events	Estimate exceedance probabilities with predictable variance-cost.	Schrödinger-bridge or Girsanov tilting; adaptive splitting/AMS; multi-index MLMC couplings.	Relative error at fixed cost; variance-decay exponent; ESS.
Benchmarking & reporting	Make results comparable and auditable across methods/costs.	Fixed data splits; structural baselines; standardized metric suite; compute/energy reporting.	Reproducible leaderboard; compute-normalized scores; energy/carbon budget.

7. Conclusion

This review set out to give researchers a compact, navigable map of recent (2020-2025) advances in stochastic modeling for biology and epidemiology, with two goals: (i) to locate and clarify *novelty* across theory, inference, learning, and numerics; and (ii) to surface *open problems* and underused methods with clear relevance to biological and epidemiological practice. We organized the field along a pipeline from Itô SDE/SPDE formulations to neural SDE identification and operator learning, reverse-time generative SDEs, and structure-preserving integrators, so that readers can

see where each contribution fits and how choices in representation, solver, and inference jointly determine scientific reliability.

Concretely, we (a) distilled an identification toolkit for drift-diffusion learning under partial observation (with PSD-constrained diffusion, monotonicity/taming, and differentiable filtering), (b) summarized operator-learning surrogates for spatial dynamics with weak-physics and ergodic/statistical alignment, and (c) reviewed solver-aware computation that maintains feasibility, stability, and long-time statistics while enabling scalable uncertainty quantification. We also charted gaps that invite immediate work in biology/epidemiology: identifiability under irregular sampling and state-dependent noise; operator learning under random forcing with well-posedness/ergodicity guarantees; discretization-aware training with memory-stable adjoints; positivity- and invariant-preserving schemes in routine pipelines; multilevel/multi-index Monte Carlo for tail-sensitive policy metrics; controlled reverse-time SDEs for assimilation and counterfactuals; and operator-level Bayesian UQ with mesh/domain generalization.

By mapping what is new, what is missing, and where methods align with biological and epidemiological objectives (time-varying transmission, spatial spread, regime shifts, and decision-oriented UQ), this review aims to shorten the path from method selection to credible results, reduce reinvention, and focus effort on the most impactful open directions.

Acknowledgments

None.

Funding

None.

Conflict of interest

There is no conflict of interest to disclose.

Author contributions

Yassine Sabbar: Software, Writing - Original draft, Writing- Review & Editing. **Kottakkaran Sooppy Nisar:** Validation, Writing- Review & Editing.

Declaration of using AI tools

The authors declare that they have not used any type of generative artificial intelligence for the writing of this manuscript, nor for the creation of images, graphics, tables, or their corresponding captions.

References

[1] S. Goldstein, B. Misra, and M. Courbage, "On intrinsic randomness of dynamical systems," J. Stat. Phys., vol. 25, no. 1, pp. 111–126, 1981. https://doi.org/10.1007/BF01008481

[2] M. Jebahi, F. Dau, J. L. Charles, and I. Iordanoff, "Multiscale modeling of complex dynamic problems: An overview and recent developments," Arch. Comput. Methods Eng., vol. 23, no. 1, pp. 101–138, 2016. https://doi.org/10.1007/s11831-014-9136-6

[3] N. Demopoulos, Foundations for a Unified Stochastic Framework: Bridging Physics. Biology, and Socio-Economic Systems, 2025.

[4] M. W. Hirsch, "The dynamical systems approach to differential equations," Bull. Am. Math. Soc., vol. 11, no. 1, pp. 1–64, 1984. https://doi.org/10.1090/S0273-0979-1984-15236-4

[5] Q. Lü and X. Zhang, Mathematical control theory for stochastic partial differential equations, vol. 101. Cham: Springer, 2021. https://doi.org/10.1007/978-3-030-82331-3

[6] M. S. Eldred, L. P. Swiler, and G. Tang, "Mixed aleatory-epistemic uncertainty quantification with stochastic expansions and optimization-based interval estimation," Reliab. Eng. Syst. Saf., vol. 96, no. 9, pp. 1092–1113, 2011. https://doi.org/10.1016/j.ress.2010.11.010

[7] R. C. Smith, Uncertainty quantification: theory, implementation, and applications. Society for Industrial and Applied Mathematics, 2024. https://doi.org/10.1137/1.9781611977844

[8] H. Schurz, "A brief review on stability investigations of numerical methods for systems of stochastic differential equations," Netw. Heterog. Media, vol. 19, no. 1, pp. 355–383, 2024. https://doi.org/10.3934/nhm.2024016

[9] N. G. Van Kampen, "Stochastic differential equations," Phys. Rep., vol. 24, no. 3, pp. 171–228, 1976. https://doi.org/10.1016/0370-1573(76)90029-6

[10] C. Salvi, M. Lemercier, and A. Gerasimovics, "Neural stochastic pdes: Resolution-invariant learning of continuous spatiotemporal dynamics," Adv. Neural Inf. Process. Syst., vol. 35, pp. 1333–1344, 2022.

[11] F. Flandoli and E. Luongo, Stochastic partial differential equations in fluid mechanics, vol. 2330. Singapore: Springer, 2023. https://doi.org/10.1007/978-981-99-0385-6

[12] E. E. Holmes, M. A. Lewis, J. E. Banks, and R. R. Veit, "Partial differential equations in ecology: Spatial interactions and population dynamics," Ecology, vol. 75, no. 1, pp. 17–29, 1994. https://doi.org/10.2307/1939378

[13] F. Arceci, M. Maurelli, D. Morale, and S. Ugolini, (2021, September). A reaction diffusion model with a stochastic boundary condition. In INdAM Workshop: Mathematical modeling and Analysis of degradation and restoration in Cultural Heritage (pp. 1-16). Singapore: Springer Nature Singapore.

[14] L. C. Evans, An introduction to stochastic differential equations, vol. 82. American Mathematical Soc., 2012.

[15] C. M. Ahn and H. E. Thompson, "Jump-diffusion processes and the term structure of interest rates," J. Finance, vol. 43, no. 1, pp. 155–174, 1988.

[16] Hewitt, K. (1970). Probabilistic approaches to discrete natural events: a review and theoretical discussion. Economic Geography, 46(sup1), 332-349. https://doi.org/10.2307/143148

[17] I. Eliazar and J. Klafter, "Lévy-driven Langevin systems: Targeted stochasticity," J. Stat. Phys., vol. 111, no. 3, pp. 739–768, 2003. https://doi.org/10.1023/A:1022894030773

[18] T. Bollerslev and V. Todorov, "Estimation of jump tails," Econometrica, vol. 79, no. 6, pp. 1727–1783, 2011. https://doi.org/10.3982/ECTA9240

[19] C. Bayer and S. Breneis, "Markovian approximations of stochastic Volterra equations with the fractional kernel," Quant. Finance, vol. 23, no. 1, pp. 53–70, 2023. https://doi.org/10.1080/14697688.2022.2139193

[20] A. Bondi and F. Flandoli, (2023). On the Kolmogorov equation associated with Volterra equations and Fractional Brownian Motion. arXiv preprint arXiv:.

[21] V. Pipiras and M. S. Taqqu, Long-range dependence and self-similarity, vol. 45. Cambridge university press, 2017. https://doi.org/10.1017/CBO9781139600347

[22] W. B. Gong, Y. Liu, V. Misra, and D. Towsley, "Self-similarity and long range dependence on the internet: A second look at the evidence, origins and implications," Comput. Netw., vol. 48, no. 3, pp. 377–399, 2005. https://doi.org/10.1016/j.comnet.2004.11.026

[23] P. C. Young, P. McKenna, and J. Bruun, "Identification of non-linear stochastic systems by state dependent parameter estimation," Int. J. Control, vol. 74, no. 18, pp. 1837–1857, 2001. https://doi.org/10.1080/00207170110089824

[24] K. Sankaran and S. P. Holmes, "Generative models: An interdisciplinary perspective," Annu. Rev. Stat. Appl., vol. 10, no. 1, pp. 325–352, 2023. https://doi.org/10.1146/annurev-statistics-033121-110134

[25] L. Pacchiardi, (2022). Statistical inference in generative models using scoring rules (Doctoral dissertation, University of Oxford).

[26] R. Cid Fernandes, L. Sodré, H. R. Schmitt, and J. R. Leao, "A probabilistic formulation for empirical population synthesis: Sampling methods and tests," Mon. Not. R. Astron. Soc., vol. 325, no. 1, pp. 60–76, 2001. https://doi.org/10.1046/j.1365-8711.2001.04366.x

[27] D. J. Albers and G. Hripcsak, "Using time-delayed mutual information to discover and interpret temporal correlation structure in complex populations," Chaos, vol. 22, no. 1, p. 013111, Mar. 2012. https://doi.org/10.1063/1.3675621

[28] P. Naveau, A. Hannart, and A. Ribes, "Statistical methods for extreme event attribution in climate science," Annu. Rev. Stat. Appl., vol. 7, no. 1, pp. 89–110, 2020. https://doi.org/10.1146/annurev-statistics-031219-041314

[29] R. J. Zhu, T. Peng, T. Cheng, X. Qu, J. Huang, D. Zhu, et al., (2025). A survey on latent reasoning. arXiv preprint arXiv:.

[30] S. Khatibisepehr, B. Huang, and S. Khare, "Design of inferential sensors in the process industry: A review of Bayesian methods," J. Process Contr., vol. 23, no. 10, pp. 1575–1596, 2013. https://doi.org/10.1016/j.jprocont.2013.05.007

[31] F. S. Richards, "A method of maximum-likelihood estimation," J. R. Stat. Soc. Series B Stat. Methodol., vol. 23, no. 2, pp. 469–475, 1961. https://doi.org/10.1111/j.2517-6161.1961.tb00430.x

[32] J. M. Wooldridge, "Applications of generalized method of moments estimation," J. Econ. Perspect., vol. 15, no. 4, pp. 87–100, 2001. https://doi.org/10.1257/jep.15.4.87

[33] G. E. Box and G. C. Tiao, Bayesian inference in statistical analysis. John Wiley & Sons, 2011.

[34] J. N. Nielsen, H. Madsen, and P. C. Young, "Parameter estimation in stochastic differential equations: An overview," Annu. Rev. Contr., vol. 24, no. 1, pp. 83–94, 2000. https://doi.org/10.1016/S1367-5788(00)00005-5

[35] R. M. Neal, (2011). MCMC using Hamiltonian dynamics. Handbook of markov chain monte carlo, 2(11), 2. 10.1201/b10905-6

[36] Septier, F., & Peters, G. W. (2015). An overview of recent advances in Monte-Carlo methods for Bayesian filtering in high-dimensional spaces. Theoretical aspects of spatial-temporal modeling, 31-61. https://doi.org/10.1007/978-4-431-55336-6_2

[37] N. Kantas, A. Doucet, S. S. Singh, and J. M. Maciejowski, (2009). An overview of sequential Monte Carlo methods for parameter estimation in general state-space models. IFAC Proceedings Volumes, 42(10), 774-785. https://doi.org/10.3182/20090706-3-FR-2004.00129

[38] D. T. Frazier and C. Drovandi, "Robust approximate Bayesian inference with synthetic likelihood," J. Comput. Graph. Stat., vol. 30, no. 4, pp. 958–976, 2021. https://doi.org/10.1080/10618600.2021.1875839

[39] J. M. Lueckmann, G. Bassetto, T. Karaletsos, and J. H. Macke, (2019, January). Likelihood-free inference with emulator networks. In Symposium on Advances in Approximate Bayesian Inference (pp. 32-53). PMLR.

[40] J. S. Chorowski, Statistics for diffusion processes with low and high-frequency observations. Germany: Humboldt Universitaet zu Berlin, 2016.

[41] M. Sorensen, (2024). Efficient estimation for ergodic diffusion processes sampled at high frequency. arXiv preprint arXiv:.

[42] Y. Aït-Sahalia and J. Jacod, "Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data," J. Econ. Lit., vol. 50, no. 4, pp. 1007–1050, 2012. https://doi.org/10.1257/jel.50.4.1007

[43] H. G. Feichtinger and K. Gröchenig, "Theory and practice of irregular sampling," in Wavelets. CRC Press, 2021, pp. 305–363. https://doi.org/10.1201/9781003210450-10

[44] F. Auteri and N. Parolini, "A mixed-basis spectral projection method," J. Comput. Phys., vol. 175, no. 1, pp. 1–23, 2002. https://doi.org/10.1006/jcph.2001.6855

[45] C. Gross and M. Iwen, "Sparse spectral methods for solving high-dimensional and multiscale elliptic PDEs," Found. Comput. Math., vol. 25 , pp. 765–811, 2025.

[46] T. Bai and P. Tahmasebi, "Hybrid geological modeling: Combining machine learning and multiple-point statistics," Comput. Geosci., vol. 142, p. 104519, 2020. https://doi.org/10.1016/j.cageo.2020.104519

[47] X. Liu, T. Xiao, S. Si, Q. Cao, S. Kumar, and C. J. Hsieh, (2019). Neural sde: Stabilizing neural ode networks with stochastic noise. arXiv preprint arXiv:.

[48] Z. Li, H. Zheng, N. Kovachki, D. Jin, H. Chen, B. Liu, et al., "Physics-informed neural operator for learning partial differential equations. ACM/IMS," J. Data Sci., vol. 1, no. 3, pp. 1–27, 2024. https://doi.org/10.6339/24-JDS1158

[49] J. L. Wu, K. Kashinath, A. Albert, D. Chirila, and H. Xiao, "Enforcing statistical constraints in generative adversarial networks for modeling chaotic dynamical systems," J. Comput. Phys., vol. 406, p. 109209, 2020. https://doi.org/10.1016/j.jcp.2019.109209

[50] J. Kwon, S. Ahn, H. Park, and M. Kwon, "Stochastic approximate inference of latent information in epidemic model: A data-driven approach," Signal Processing, vol. 232, p. 109919, 2025. https://doi.org/10.1016/j.sigpro.2025.109919

[51] A. Ruszczyński and A. Shapiro, "Stochastic programming models," Handb. Oper. Res. Manag. Sci., vol. 10, pp. 1–64, 2003. https://doi.org/10.1016/S0927-0507(03)10001-1

[52] A. Shapiro, D. Dentcheva, and A. Ruszczynski, Lectures on stochastic programming: modeling and theory. Society for Industrial and Applied Mathematics, 2021. https://doi.org/10.1137/1.9781611976595

[53] A. Cohen, W. Dahmen, R. A. DeVore, and A. Kunoth, "Multiscale and High-Dimensional Problems," Oberwolfach Reports, vol. 14, no. 1, pp. 1001–1051, 2018. https://doi.org/10.4171/owr/2017/17

[54] Y. Zhang, G. Wang, F. Su, and Z. Qiao, "The theory of response analysis of fuzzy stochastic dynamical systems with a single degree of freedom," Earthquake Eng. Struct. Dynam., vol. 25, no. 3, pp. 235–251, 1996. https://doi.org/10.1002/(SICI)1096-9845(199603)25:3<235::AID-EQE552>3.0.CO;2-3

[55] K. R. Sreenivasan, "Fluid turbulence," Rev. Mod. Phys., vol. 71, no. 2, pp. S383–S395, 1999. https://doi.org/10.1103/RevModPhys.71.S383

[56] F. Noé, A. Tkatchenko, K. R. Müller, and C. Clementi, "Machine learning for molecular simulation," Annu. Rev. Phys. Chem., vol. 71, no. 1, pp. 361–390, Apr. 20 2020. https://doi.org/10.1146/annurev-physchem-042018-052331

[57] R. C. Malone, R. D. Smith, M. E. Maltrud, and M. W. Hecht, "Eddy-resolving ocean modeling," Los Alamos Sci., vol. 28, pp. 223–231, 2003.

[58] M. Aoki, Optimization of stochastic systems: topics in discrete-time systems. Elsevier, 2016.

[59] Y. Huang, X. F. Zha, J. Lee, and C. Liu, "Discriminant diffusion maps analysis: A robust manifold learner for dimensionality reduction and its applications in machine condition monitoring and fault diagnosis," Mech. Syst. Signal Process., vol. 34, no. 1-2, pp. 277–297, 2013. https://doi.org/10.1016/j.ymssp.2012.04.021

[60] Abdulle, A., & Bai, Y. (2014). Reduced-order modelling numerical homogenization. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 372(2021), 20130388. https://doi.org/10.1098/rsta.2013.0388

[61] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," Nonlinear Dyn., vol. 41, no. 1, pp. 309–325, 2005. https://doi.org/10.1007/s11071-005-2824-x

[62] C. Kelly and G. J. Lord, "Adaptive Euler methods for stochastic systems with non-globally Lipschitz coefficients," Numer. Algorithms, vol. 89, no. 2, pp. 721–747, 2022. https://doi.org/10.1007/s11075-021-01131-8

[63] X. Chen, G. Dos Reis, and W. Stockinger, "Wellposedness, exponential ergodicity and numerical approximation of fully super-linear McKean-Vlasov SDEs and associated particle systems," Electron. J. Probab., vol. 30, none, pp. 1–50, 2025. https://doi.org/10.1214/25-EJP1278

[64] Y. Hu, "Semi-implicit Euler-Maruyama scheme for stiff stochastic equations," in Stochastic Analysis and Related Topics V: The Silivri Workshop, 1994. Boston, MA: Birkhäuser Boston, 1996, pp. 183–202. https://doi.org/10.1007/978-1-4612-2450-1_9

[65] T. Sands, "Treatise on analytic nonlinear optimal guidance and control amplification of strictly analytic (non-numerical) methods," Front. Robot. AI, vol. 9, p. 884669, Oct. 4 2022. https://doi.org/10.3389/frobt.2022.884669

[66] U. M. Ascher, S. J. Ruuth, and B. T. Wetton, "Implicit-explicit methods for time-dependent partial differential equations," SIAM J. Numer. Anal., vol. 32, no. 3, pp. 797–823, 1995. https://doi.org/10.1137/0732037

[67] M. Hochbruck and A. Ostermann, "Exponential integrators," Acta Numer., vol. 19, pp. 209–286, 2010. https://doi.org/10.1017/S0962492910000048

[68] T. Y. A. Averina and K. A. Rybakov, "Rosenbrock-type methods for solving stochastic differential equations," Numer. Anal. Appl., vol. 17, no. 2, pp. 99–115, 2024. https://doi.org/10.1134/S1995423924020010

[69] S. Blanes, A. Iserles, and S. Macnamara, "Positivity-preserving methods for ordinary differential equations," Model. Math. Anal. Numer., vol. 56, no. 6, pp. 1843–1870, 2022. https://doi.org/10.1051/m2an/2022042

[70] Y. Sabbar and S. F. Aldosary, "A Novel Stochastic SVIR Model Capturing Transmission Variability Through Mean-Reverting Processes and Stationary Reproduction Thresholds," Mathematics, vol. 13, no. 13, p. 2097, 2025. https://doi.org/10.3390/math13132097

[71] A. Khan, Y. Sabbar, and A. Din, "Stochastic modeling of the Monkeypox 2022 epidemic with cross-infection hypothesis in a highly disturbed environment," Math. Biosci. Eng., vol. 19, no. 12, pp. 13560–13581, Sep. 15 2022. https://doi.org/10.3934/mbe.2022633

[72] N. Privault and L. Wang, "Stochastic SIR Lévy jump model with heavy-tailed increments," J. Nonlinear Sci., vol. 31, no. 1, p. 15, 2021. https://doi.org/10.1007/s00332-020-09670-5

[73] Y. Sabbar, (2021). Mathematical Analysis of Some Stochastic Infectious Disease Models with White Noises and Lévy Jumps (Doctoral dissertation, Université Sidi Mohamed Ben Abdellah de Fès (Maroc)).

[74] Y. Sabbar, "Asymptotic extinction and persistence of a perturbed epidemic model with different intervention measures and standard lévy jumps," Bulletin of Biomathematics, vol. 1, no. 1, pp. 58–77, 2023. https://doi.org/10.59292/bulletinbiomath.2023004

[75] A. Alkhazzan, Y. Guoqiang, W. Zhou, S. M. A. Shah, and Y. Sabbar, "Stochastic modeling of worm propagation in PLC-based industrial systems with general incidence and Lévy noise," Chaos Solitons Fractals, vol. 202, p. 117563, 2026. https://doi.org/10.1016/j.chaos.2025.117563

[76] Y. Sabbar, "Refining extinction criteria in a complex multi-stage epidemic system with non-Gaussian Lévy noise," Commun. Nonlinear Sci. Numer. Simul., vol. 149, p. 108911, 2025. https://doi.org/10.1016/j.cnsns.2025.108911

[77] M. C. Mariani and I. Florescu, Quantitative finance. John Wiley and Sons, 2019. https://doi.org/10.1002/9781118630006

[78] K. C. Kapur and M. G. Pecht, Reliability engineering. John Wiley and Sons, 2014. https://doi.org/10.1002/9781118841716

[79] B. M. Althouse, E. A. Wenger, J. C. Miller, S. V. Scarpino, A. Allard, L. Hébert-Dufresne, et al., "Superspreading events in the transmission dynamics of SARS-CoV-2: Opportunities for interventions and control," PLoS Biol., vol. 18, no. 11, p. e3000897, Nov. 12 2020. https://doi.org/10.1371/journal.pbio.3000897

[80] B. Pei, L. Feng, Y. Li, and Y. Xu, "Non-markovian dynamics: The memory-dependent probability density evolution equations," Nonlinear Dyn., vol. 113, no. 11, pp. 1–19, 2025. https://doi.org/10.1007/s11071-025-11078-3

[81] B. P. Moghaddam, A. Babaei, A. Dabiri, and A. Galhano, "Fractional stochastic partial differential equations: Numerical advances and practical applications—A state of the art review," Symmetry (Basel), vol. 16, no. 5, p. 563, 2024. https://doi.org/10.3390/sym16050563

[82] C. Li, D. Qian, and Y. Chen, "On Riemann-Liouville and caputo derivatives," Discrete Dyn. Nat. Soc., vol. 2011, no. 1, p. 562494, 2011. https://doi.org/10.1155/2011/562494

[83] R. Metzler, W. G. Glöckle, and T. F. Nonnenmacher, "Fractional model equation for anomalous diffusion," Physica A, vol. 211, no. 1, pp. 13–24, 1994. https://doi.org/10.1016/0378-4371(94)90064-7

[84] K. Adolfsson, M. Enelund, and P. Olsson, "On the fractional order model of viscoelasticity," Mech. Time-Depend. Mater., vol. 9, no. 1, pp. 15–34, 2005. https://doi.org/10.1007/s11043-005-3442-1

[85] C. H. Lee, C. H. Juan, H. H. Chen, J. P. Hong, T. W. Liao, I. French, et al., "Long-Range Temporal Correlations in Electroencephalography for Parkinson's Disease Progression," Mov. Disord., vol. 40, no. 2, pp. 266–275, Feb. 2025. https://doi.org/10.1002/mds.30074

[86] Voronenko, V. I., Kovalov, B. L., Kharchenko, M. O., Hrytsenko, P. V., & Omelianenko, V. A. (2024). The Development of the digital transformation of socio-economic and ecological systems.

[87] L. Beghin, F. Mainardi, and R. Garrappa, Eds. Nonlocal and fractional operators. Cham, Switzerland: Springer, 2021. https://doi.org/10.1007/978-3-030-69236-0

[88] J. Ren and Y. Wang, "Overview of object detection algorithms using convolutional neural networks," Journal of Computer and Communications, vol. 10, no. 1, pp. 115–132, 2022.

[89] Z. Lyu, T. Yu, F. Pan, Y. Zhang, J. Luo, D. Zhang, et al., "A survey of model compression strategies for object detection," Multimedia Tools Appl., vol. 83, no. 16, pp. 48165–48236, 2024. https://doi.org/10.1007/s11042-023-17192-x

[90] R. Motorga, V. Mureșan, M. L. Ungureșan, M. Abrudean, H. Vălean, and I. Clitan, "Artificial intelligence in fractional-order systems approximation with high performances: Application in modelling of an isotopic separation process," Mathematics, vol. 10, no. 9, p. 1459, 2022. https://doi.org/10.3390/math10091459

[91] M. Hosseini, M. Horton, H. Paneliya, U. Kallakuri, H. Homayoun, and T. Mohsenin, (2019, June). On the complexity reduction of dense layers from o (n2) to o (nlogn) with cyclic sparsely connected layers. In Proceedings of the 56th Annual Design Automation Conference 2019 (pp. 1-6). https://doi.org/10.1145/3316781.3317873

[92] I. Pan, L. R. Mason, and O. K. Matar, "Data-centric Engineering: Integrating simulation, machine learning and statistics. Challenges and opportunities," Chem. Eng. Sci., vol. 249, p. 117271, 2022. https://doi.org/10.1016/j.ces.2021.117271

[93] M. Shen and C. Cheng, (2025). Neural SDEs as a Unified Approach to Continuous-Domain Sequence Modeling. arXiv preprint arXiv:.

[94] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor, "Sample, quantize, and encode: Timely estimation over noisy channels," IEEE Trans. Commun., vol. 69, no. 10, pp. 6485–6499, 2021. https://doi.org/10.1109/TCOMM.2021.3092413

[95] L. Yang, S. Liu, T. Meng, and S. J. Osher, "In-context operator learning with data prompts for differential equation problems," Proc. Natl. Acad. Sci. USA, vol. 120, no. 39, p. e2310142120, Sep. 26 2023. https://doi.org/10.1073/pnas.2310142120

[96] N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, et al., "Neural operator: Learning maps between function spaces with applications to pdes," J. Mach. Learn. Res., vol. 24, no. 89, pp. 1–97, 2023.

[97] H. Cao, C. Tan, Z. Gao, Y. Xu, G. Chen, P. A. Heng, et al., "A survey on generative diffusion models," IEEE Trans. Knowl. Data Eng., vol. 36, no. 7, pp. 2814–2830, 2024. https://doi.org/10.1109/TKDE.2024.3361474

[98] P. Kumar, "Diffusion Models and Generative Artificial Intelligence: Frameworks, Applications and Challenges," Arch. Comput. Methods Eng., vol. 32, no. 7, pp. 1–44, 2025. https://doi.org/10.1007/s11831-025-10266-z

[99] P. Johri, C. F. Aquadro, M. Beaumont, B. Charlesworth, L. Excoffier, A. Eyre-Walker, et al., "Recommendations for improving statistical inference in population genomics," PLoS Biol., vol. 20, no. 5, p. e3001669, May 31 2022. https://doi.org/10.1371/journal.pbio.3001669

[100] Z. Hu, H. Wang, and Q. Zhou, "A MCMC method based on surrogate model and Gaussian process parameterization for infinite Bayesian PDE inversion," J. Comput. Phys., vol. 507, p. 112970, 2024. https://doi.org/10.1016/j.jcp.2024.112970

[101] J. Donnelly, S. Abolfathi, J. Pearson, O. Chatrabgoun, and A. Daneshkhah, "Gaussian process emulation of spatio-temporal outputs of a 2D inland flood model," Water Res., vol. 225, p. 119100, Oct. 15 2022. https://doi.org/10.1016/j.watres.2022.119100

[102] D. V. Patel, J. Lee, M. W. Farthing, P. K. Kitanidis, and E. F. Darve, (2024). Multi-fidelity Hamiltonian Monte Carlo. arXiv preprint arXiv:.

[103] J. Zhang, "Modern Monte Carlo methods for efficient uncertainty quantification and propagation: A survey," Wiley Interdiscip. Rev. Comput. Stat., vol. 13, no. 5, p. e1539, 2021. https://doi.org/10.1002/wics.1539

[104] M. Nourali, "Comparison of likelihood-free inference approach and a formal Bayesian method in parameter uncertainty assessment: Case study with a single-event rainfall-runoff model," J. Hydrol. Eng., vol. 26, no. 3, p. 05020049, 2021. https://doi.org/10.1061/(ASCE)HE.1943-5584.0002048

[105] D. T. Frazier, D. J. Nott, C. Drovandi, and R. Kohn, "Bayesian inference using synthetic likelihood: Asymptotics and adjustments," J. Am. Stat. Assoc., vol. 118, no. 544, pp. 2821–2832, 2023. https://doi.org/10.1080/01621459.2022.2086132

[106] T. Granger, T. M. Michelitsch, M. Bestehorn, A. P. Riascos, and B. A. Collet, "Stochastic compartment model with mortality and its application to epidemic spreading in complex networks," Entropy (Basel), vol. 26, no. 5, p. 362, Apr. 25 2024. https://doi.org/10.3390/e26050362

[107] Y. Sabbar, "Exploring threshold dynamics of a behavioral epidemic model featuring two susceptible classes and second-order jump-diffusion," Chaos Solitons Fractals, vol. 186, p. 115216, 2024. https://doi.org/10.1016/j.chaos.2024.115216

[108] K. S. Nisar and Y. Sabbar, "Long-run analysis of a perturbed HIV/AIDS model with antiretroviral therapy and heavy-tailed increments performed by tempered stable Lévy jumps," Alex. Eng. J., vol. 78, pp. 498–516, 2023. https://doi.org/10.1016/j.aej.2023.07.053

[109] D. Kiouach and Y. Sabbar, (2019, April). Threshold analysis of the stochastic Hepatitis B epidemic model with successful vaccination and Levy jumps. In 2019 4th World Conference on Complex Systems (WCCS) (pp. 1-6). IEEE.

[110] D. Kiouach and Y. Sabbar, (2019, February). Modeling the impact of media intervention on controlling the diseases with stochastic perturbations. In AIP Conference Proceedings (Vol. 2074, No. 1, p. 020026). AIP Publishing LLC. https://doi.org/10.1063/1.5090643

[111] D. Wiwatanapataphee, Y. H. Wu, W. Sawangtong, P. Sawangtong, and W. Australia, "Modeling anomalous diffusion and volatility in the Australian national electricity market using a space-fractional Black-Scholes framework," AIMS Math., vol. 10, no. 5, pp. 12388–12420, 2025. https://doi.org/10.3934/math.2025560

[112] K. A. Newhall, "A Primer on Stochastic Partial Differential Equations with Spatially Correlated Noise," Annu. Rev. Condens. Matter Phys., vol. 16 , pp.195–208, 2025.

[113] E. M. B. Bouhadjar, Backward SDEs and Applications to Optimal Control Problems (Doctoral dissertation, Université Mohamed Khider (Biskra-Algérie)), 2024.

[114] U. Sirisha, S. P. Praveen, P. N. Srinivasu, P. Barsocchi, and A. K. Bhoi, "Statistical analysis of design aspects of various YOLO-based deep learning models for object detection," International Journal of Computational Intelligence Systems, vol. 16, no. 1, p. 126, 2023. https://doi.org/10.1007/s44196-023-00302-w

[115] M. Kaveh and M. S. Mesgari, "Application of meta-heuristic algorithms for training neural networks and deep learning architectures: A comprehensive review," Neural Process. Lett., vol. 55, no. 4, pp. 1–104, Oct. 31 2022.

[116] N. Gaby, X. Ye, and H. Zhou, "Neural control of parametric solutions for high-dimensional evolution pdes," SIAM J. Sci. Comput., vol. 46, no. 2, pp. C155–C185, 2024. https://doi.org/10.1137/23M1549870

[117] Z. Gao, L. Yan, and T. Zhou, "Adaptive operator learning for infinite-dimensional Bayesian inverse problems," SIAM/ASA J. Uncertainty Quantif., vol. 12, no. 4, pp. 1389–1423, 2024. https://doi.org/10.1137/24M1643815

[118] K. Suresh, C. Severn, and D. Ghosh, "Survival prediction models: An introduction to discrete-time modeling," BMC Med. Res. Methodol., vol. 22, no. 1, p. 207, Jul. 26 2022. https://doi.org/10.1186/s12874-022-01679-6

[119] A. Kaneko, "Multi-stage Euler-Maruyama methods for backward stochastic differential equations driven by continuous-time Markov chains," Jpn. J. Ind. Appl. Math., vol. 41, no. 2, pp. 1223–1276, 2024. https://doi.org/10.1007/s13160-024-00649-3

[120] J. Cai, R. J. B. Goudie, C. Starr, and B. D. M. Tom, "Dynamic factor analysis with dependent Gaussian processes for high-dimensional gene expression trajectories," Biometrics, vol. 80, no. 4,ujae131, Oct. 3 2024. https://doi.org/10.1093/biomtc/ujae131

[121] H. Hoel and S. Ragunathan, "Higher-order adaptive methods for exit times of Itô diffusions," IMA J. Numer. Anal., vol. 44, no. 5, pp. 2821–2863, 2024. https://doi.org/10.1093/imanum/drad077

[122] N. J. Higham, "Cholesky factorization," Wiley Interdiscip. Rev. Comput. Stat., vol. 1, no. 2, pp. 251–254, 2009. https://doi.org/10.1002/wics.18

[123] M. Gonzalez, N. Fernandez Pinto, T. Tran, H. Hajri, and N. Masmoudi, "Seeds: Exponential sde solvers for fast high-quality sampling from diffusion models," Adv. Neural Inf. Process. Syst., vol. 36, pp. 68061–68120, 2023.

[124] D. N. Tanyu, J. Ning, T. Freudenberg, N. Heilenkötter, A. Rademacher, U. Iben, et al., "Deep learning methods for partial differential equations and related parameter identification problems," Inverse Probl., vol. 39, no. 10, p. 103001, 2023. https://doi.org/10.1088/1361-6420/ace9d4

[125] W. Bao, S. Feng, and K. Zhang, (2024). Application of Kalman Filter in Stochastic Differential Equations. arXiv preprint arXiv:.

[126] G. Boaretto and M. P. Laurini, "DSGE Estimation Using Generalized Empirical Likelihood and Generalized Minimum Contrast," Entropy (Basel), vol. 27, no. 2, p. 141, Jan. 30 2025. https://doi.org/10.3390/e27020141

[127] W. Davis and B. Buffett, "Estimation of drift and diffusion functions from unevenly sampled time-series data," Phys. Rev. E, vol. 106, 1-1, p. 014140, Jul. 2022. https://doi.org/10.1103/PhysRevE.106.014140

[128] P. R. Hansen and G. Horel, (2009). Quadratic variation by Markov chains. Univ. of Aarhus Dept. of Economics Research Paper, (2009-13). https://doi.org/10.2139/ssrn.1367519

[129] G. H. Choe, "Girsanov's theorem," in Stochastic Analysis for Finance with Simulations. Cham: Springer International Publishing, 2016, pp. 137–145. https://doi.org/10.1007/978-3-319-25589-7_8

[130] A. L. Lewis, (2001). A simple option formula for general jump-diffusion and other exponential Lévy processes. Available at SSRN 282110. https://doi.org/10.2139/ssrn.282110

[131] L. J. Powers, J. Nešlehová, and D. A. Stephens, "Pricing American Options in an Infinite Activity Lévy Market: Monte Carlo and Deterministic Approaches Using a Diffusion Approximation," in Numerical Methods in Finance: Bordeaux, June 2010. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 291–321. https://doi.org/10.1007/978-3-642-25746-9_9

[132] V. Guan, J. Janssen, H. Rahmani, A. Warren, S. Zhang, E. Robeva, et al., (2024). Identifying drift, diffusion, and causal structure from temporal snapshots. arXiv preprint arXiv:.

[133] Y. Chen and M. J. Wainwright, (2015). Fast low-rank estimation by projected gradient descent: General statistical and algorithmic guarantees. arXiv preprint arXiv:.

[134] S. N. Sharma and H. G. Patel, "The Fokker-Planck equation," in Stochastic Control. Rijeka: IntechOpen, 2010, pp. 1–20.

[135] L. Borcea, T. Hohage, and B. Kaltenbacher, "Computational Inverse Problems for Partial Differential Equations," Oberwolfach Reports, vol. 14, no. 2, pp. 1463–1549, 2018. https://doi.org/10.4171/owr/2017/24

[136] Z. Zou, X. Meng, A. F. Psaros, and G. E. Karniadakis, "NeuralUQ: A comprehensive library for uncertainty quantification in neural differential equations and operators," SIAM Rev., vol. 66, no. 1, pp. 161–190, 2024. https://doi.org/10.1137/22M1518189

[137] P. Bera and S. Bhanja, (2025). Quantification of Uncertainties in Probabilistic Deep Neural Network by Implementing Boosting of Variational Inference. arXiv preprint arXiv:.

[138] X. Chen, (2024). Enhancing Estimation and Uncertainty Quantification in Stochastic Optimization: Importance Sampling and Bootstrap Resampling (Doctoral dissertation, University of California, Davis).

[139] X. Feng, H. Sha, Y. Zhang, Y. Su, S. Liu, Y. Jiang, et al., "Reliable deep learning in anomalous diffusion against out-of-distribution dynamics," Nat. Comput. Sci., vol. 4, no. 10, pp. 761–772, Oct. 2024. https://doi.org/10.1038/s43588-024-00703-7

[140] M. Fuhrman, F. Masiero, and G. Tessitore, "Stochastic equations with delay: Optimal control via BSDEs and regular solutions of Hamilton-Jacobi-Bellman equations," SIAM J. Contr. Optim., vol. 48, no. 7, pp. 4624–4651, 2010. https://doi.org/10.1137/080730354

[141] S. Lototsky and B. Rozovskii, "Stochastic differential equations: a Wiener chaos approach," in From stochastic calculus to mathematical finance: the Shiryaev festschrift. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 433–506. https://doi.org/10.1007/978-3-540-30788-4_23

[142] S. V. Lototsky and B. L. Rozovsky, Stochastic partial differential equations, vol. 11. Cham: Springer, 2017. https://doi.org/10.1007/978-3-319-58647-2

[143] M. Liu-Schiaffini, J. Berner, B. Bonev, T. Kurth, K. Azizzadenesheli, and A. Anandkumar, (2024). Neural operators with localized integral and differential kernels. arXiv preprint arXiv:.

[144] O. R. Dunbar, N. H. Nelsen, and M. Mutic, "Hyperparameter optimization for randomized algorithms: A case study on random features," Stat. Comput., vol. 35, no. 3, pp. 1–28, 2025. https://doi.org/10.1007/s11222-025-10587-w

[145] S. Li, S. U. Khan, K. Yu, E. A. Ismail, and F. A. Awwad, "Stability Analysis of Spectral Technique for Three-Connected Fractional Stochastic Integro-Delay Differential Systems," Int. J. Theor. Phys., vol. 64, no. 4, pp. 1–29, 2025. https://doi.org/10.1007/s10773-025-05969-9

[146] S. Chen, (2025). Knowledge-Guided Machine Learning Methods and Applications to Scientific Data Simulation and Upscaling (Doctoral dissertation, University of Pittsburgh).

[147] N. Chen, Stochastic methods for modeling and predicting complex dynamical systems. Berlin/Heidelberg, Germany: Springer International Publishing, 2023. https://doi.org/10.1007/978-3-031-22249-8

[148] G. Picci, "Stochastic Noises, Observation, Identification and Realization with," in Encyclopedia of Complexity and Systems Science. New York, NY: Springer, 2009, pp. 8749–8766. https://doi.org/10.1007/978-0-387-30440-3_525

[149] M. Grigoriu, "Evaluation of Karhunen-Loève, spectral, and sampling representations for stochastic processes," J. Eng. Mech., vol. 132, no. 2, pp. 179–189, 2006. https://doi.org/10.1061/(ASCE)0733-9399(2006)132:2(179)

[150] B. Haasdonk, K. Urban, and B. Wieland, "Reduced basis methods for parameterized partial differential equations with stochastic influences using the Karhunen-Loeeve expansion," SIAM/ASA J. Uncertainty Quantif., vol. 1, no. 1, pp. 79–105, 2013. https://doi.org/10.1137/120876745

[151] R. Kupferman, G. A. Pavliotis, and A. M. Stuart, "Itô versus Stratonovich white-noise limits for systems with inertia and colored multiplicative noise," Phys. Rev. E Stat. Nonlin. Soft Matter Phys., vol. 70, 3 Pt 2, p. 036120, Sep. 2004. https://doi.org/10.1103/PhysRevE.70.036120

[152] J. Aaronson, An introduction to infinite ergodic theory (No. 50). American Mathematical Soc., 1997. https://doi.org/10.1090/surv/050

[153] B. Després, M. Dumbser, J. R. Kamm, and M. Torrilhon, "High-Resolution Mathematical and Numerical Analysis of Involution-Constrained PDEs," Oberwolfach Reports, vol. 10, no. 3, pp. 2691–2747, 2014. https://doi.org/10.4171/owr/2013/47

[154] A. Abdulle and G. R. de Souza, "Explicit stabilized multirate method for stiff stochastic differential equations," SIAM J. Sci. Comput., vol. 44, no. 4, pp. A1859–A1883, 2022. https://doi.org/10.1137/21M1439018

[155] K. Koval and R. Nicholson, "Non-intrusive optimal experimental design for large-scale nonlinear Bayesian inverse problems using a Bayesian approximation error approach: K.

Koval, R. Nicholson," J. Sci. Comput., vol. 104, no. 3, p. 98, 2025. https://doi.org/10.1007/s10915-025-03008-7

[156] J. Watson and J. Peters, (2021, May). Advancing trajectory optimization with approximate inference: Exploration, covariance control and adaptive risk. In 2021 American Control Conference (ACC) (pp. 1231-1236). IEEE. https://doi.org/10.23919/ACC50511.2021.9482657

[157] H. U. Molla and J. Qiu, "Numerical approximations of coupled forward-backward SPDEs," Stochastic Anal. Appl., vol. 41, no. 2, pp. 291–326, 2023. https://doi.org/10.1080/07362994.2021.2011318

[158] Z. Hao, S. Liu, Y. Zhang, C. Ying, Y. Feng, H. Su, et al., (2022). Physics-informed machine learning: A survey on problems, methods and applications. arXiv preprint arXiv:.

[159] P. Glasserman, Gradient estimation via perturbation analysis, vol. 116. Springer Science & Business Media, 1991.

[160] A. Taghvaei, (2024, December). Time-Reversal of Stochastic Maximum Principle. In 2024 IEEE 63rd Conference on Decision and Control (CDC) (pp. 2597-2602). IEEE. https://doi.org/10.1109/CDC56724.2024.10886762

[161] H. T. Nguyen, J. B. Andersen, G. F. Pedersen, P. Kyritsi, and P. C. Eggers, "Time reversal in wireless communications: A measurement-based investigation," IEEE Trans. Wirel. Commun., vol. 5, no. 8, pp. 2242–2252, 2006. https://doi.org/10.1109/TWC.2006.1687740

[162] B. E. Anderson, R. A. Guyer, T. J. Ulrich, and P. A. Johnson, "Time reversal of continuous-wave, steady-state signals in elastic media," Appl. Phys. Lett., vol. 94, no. 11, p. 111908, 2009. https://doi.org/10.1063/1.3097811

[163] A. Li, W. Fang, H. Zhao, L. Lu, G. Yang, and M. Xu, (2025). MaRS: A Fast Sampler for Mean Reverting Diffusion based on ODE and SDE Solvers. arXiv preprint arXiv:.

[164] A. Solin, E. Tamir, and P. Verma, "Scalable inference in SDEs by direct matching of the Fokker-Planck-Kolmogorov equation," Adv. Neural Inf. Process. Syst., vol. 34, pp. 417–429, 2021.

[165] S. Xue, M. Yi, W. Luo, S. Zhang, J. Sun, Z. Li, et al., "Sa-solver: Stochastic adams solver for fast sampling of diffusion models," Adv. Neural Inf. Process. Syst., vol. 36, pp. 77632–77674, 2023.

[166] E. Dussaud, W. W. Symes, P. Williamson, L. Lemaistre, P. Singer, B. Denel, et al., "Computational strategies for reverse-time migration," SEG Tech. Prog. Exp. Abstr., vol. 2008, pp. 2267–2271, 2008. (Society of Exploration Geophysicists.)

[167] S. Chen, S. Chewi, H. Lee, Y. Li, J. Lu, and A. Salim, "The probability flow ode is provably fast," Adv. Neural Inf. Process. Syst., vol. 36, pp. 68552–68575, 2023.

[168] M. M. Chertkov, "Mixing artificial and natural intelligence: From statistical mechanics to AI and back to turbulence," J. Phys. A Math. Theor., vol. 57, no. 33, p. 333001, 2024. https://doi.org/10.1088/1751-8121/ad67bb

[169] G. Fabiani, (2024). Machine learning-based Modelling and Numerical Analysis of the Emergent Dynamics of Complex Systems (Doctoral dissertation, University of Naples Federico II).

[170] A. Agrawal and C. McComb, (2025). Adaptive Multi-Fidelity Reinforcement Learning for Variance Reduction in Engineering Design Optimization. arXiv preprint arXiv:.

[171] P. Sanchez and S. A. Tsaftaris, (2022). Diffusion causal models for counterfactual estimation. arXiv preprint arXiv:.

[172] H. Hassani, R. Razavi-Far, M. Saif, and L. Lin, (2024). Towards Sample-Efficiency and Generalization of Transfer and Inverse Reinforcement Learning: A Comprehensive Literature Review. arXiv preprint arXiv:.

[173] P. Behnoudfar and N. Chen, (2025). RL-DAUNCE: Reinforcement Learning-Driven Data Assimilation with Uncertainty-Aware Constrained Ensembles. arXiv preprint arXiv:. https://doi.org/10.2139/ssrn.5259995

[174] W. Rüemelin, "Numerical treatment of stochastic differential equations," SIAM J. Numer. Anal., vol. 19, no. 3, pp. 604–613, 1982. https://doi.org/10.1137/0719041

[175] W. Fang and M. B. Giles, (2016, August). Adaptive Euler-Maruyama method for SDEs with non-globally Lipschitz drift. In International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing (pp. 217-234). Cham: Springer International Publishing.

[176] M. Bossy and K. Martínez, (2024). Strong convergence of the exponential Euler scheme for SDEs with superlinear growth coefficients and one-sided Lipschitz drift. arXiv preprint arXiv:.

[177] A. Lang, "Almost sure convergence of a Galerkin approximation for SPDEs of Zakai type driven by square integrable martingales," J. Comput. Appl. Math., vol. 236, no. 7, pp. 1724–1732, 2012. https://doi.org/10.1016/j.cam.2011.10.003

[178] Z. Jia and C. Li, "Almost sure exponential stability of uncertain stochastic Hopfield neural networks based on subadditive measures," Mathematics, vol. 11, no. 14, p. 3110, 2023. https://doi.org/10.3390/math11143110

[179] M. Sharp, (2024). Quantitative Preservations of Ulam Stability-type Estimates. arXiv preprint arXiv:.

[180] C. E. Bréhier, "Approximation of the invariant distribution for a class of ergodic SDEs with one-sided Lipschitz continuous drift coefficient using an explicit tamed Euler scheme," ESAIM Probab. Stat., vol. 27, pp. 841–866, 2023. https://doi.org/10.1051/ps/2023017

[181] G. Rohith and N. K. Sinha, "Routes to chaos in the post-stall dynamics of higher-dimensional aircraft model," Nonlinear Dyn., vol. 100, no. 2, pp. 1705–1724, 2020. https://doi.org/10.1007/s11071-020-05604-8

[182] R. D'Ambrosio, "Numerical Methods for Stochastic Differential Equations," in Numerical Approximation of Ordinary Differential Problems: From Deterministic to Stochastic Numerical Methods. Cham: Springer Nature Switzerland, 2023, pp. 291–363. https://doi.org/10.1007/978-3-031-31343-1_9

[183] F. Djeumou, T. J. Lew, N. Ding, M. Thompson, M. Suminaka, M. Greiff, et al., (2024, November). One model to drift them all: Physics-informed conditional diffusion model for driving at the limits. In 8th Annual Conference on Robot Learning.

[184] C. K. Wai, L. K. Choon, and R. Idris, "Adaptive time-stepping for Runge-Kutta method for ordinary differential equations," Universiti Malaysia Terengganu Journal of Undergraduate Research, vol. 3, no. 1, pp. 25–36, 2021. https://doi.org/10.46754/umtjur.v3i1.189

[185] Z. Liu and Z. Liu, (2024). Unique Ergodicity of Implicit Methods for Monotone SDEs and SPDEs driven by Nondegenerate Multiplicative Noise. arXiv preprint arXiv:.

[186] J. Kuntz, P. Thomas, G. B. Stan, and M. Barahona, "Stationary distributions of continuous-time Markov chains: A review of theory and truncation-based approximations," SIAM Rev., vol. 63, no. 1, pp. 3–64, 2021. https://doi.org/10.1137/19M1289625

[187] A. Ni, "Ergodic and foliated kernel-differentiation method for linear responses of random systems," J. Nonlinear Sci., vol. 35, no. 5, p. 90, 2025. https://doi.org/10.1007/s00332-025-10186-z

[188] M. Perman, J. Pitman, and M. Yor, "Size-biased sampling of Poisson point processes and excursions," Probab. Theory Relat. Fields, vol. 92, no. 1, pp. 21–39, 1992. https://doi.org/10.1007/BF01205234

[189] D. Talay and L. Tubaro, "Expansion of the global error for numerical schemes solving stochastic differential equations," Stochastic Anal. Appl., vol. 8, no. 4, pp. 483–509, 1990. https://doi.org/10.1080/07362999008809220

[190] J. S. Lamb and J. A. Roberts, "Time-reversal symmetry in dynamical systems: A survey," Physica D, vol. 112, no. 1-2, pp. 1–39, 1998. https://doi.org/10.1016/S0167-2789(97)00199-1

[191] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," Nat. Rev. Phys., vol. 3, no. 6, pp. 422–440, 2021. https://doi.org/10.1038/s42254-021-00314-5

[192] R. Dascaliuc, C. Foias, and M. S. Jolly, "Estimates on enstrophy, palinstrophy, and invariant measures for 2-D turbulence," J. Differ. Equ., vol. 248, no. 4, pp. 792–819, 2010. https://doi.org/10.1016/j.jde.2009.11.020

[193] D. Goodair and D. Crisan, Stochastic Calculus in Infinite Dimensions and SPDEs. Springer Nature, 2024. https://doi.org/10.1007/978-3-031-69586-5

[194] F. Y. Wang, "Exponential contraction in Wasserstein distances for diffusion semigroups with negative curvature," Potential Anal., vol. 53, no. 3, pp. 1123–1144, 2020. https://doi.org/10.1007/s11118-019-09800-z

[195] M. B. Giles, "Multilevel monte carlo methods," Acta Numer., vol. 24, pp. 259–328, 2015. https://doi.org/10.1017/S096249291500001X

[196] F. Cornalba and J. Fischer, "Multilevel Monte Carlo methods for the Dean-Kawasaki equation from fluctuating hydrodynamics," SIAM J. Numer. Anal., vol. 63, no. 1, pp. 262–287, 2025. https://doi.org/10.1137/23M1617345

[197] W. Fang and M. B. Giles, "Adaptive Euler-Maruyama method for SDEs with nonglobally Lipschitz drift," Ann. Appl. Probab., vol. 30, no. 2, pp. 526–560, 2020. https://doi.org/10.1214/19-AAP1507

[198] C. Bayer, P. K. Friz, S. Riedel, and J. Schoenmakers, "From rough path estimates to multilevel Monte Carlo," SIAM J. Numer. Anal., vol. 54, no. 3, pp. 1449–1483, 2016. https://doi.org/10.1137/140995209

[199] P. Mycek, (2024). Hierarchical methods for deterministic and stochastic partial differential equations (Doctoral dissertation, Université de Bordeaux).

[200] Y. Bai, Z. Huang, H. Lam, and D. Zhao, "Overconservativeness of variance-based efficiency criteria and probabilistic efficiency in rare-event simulation," Manage. Sci., vol. 70, no. 10, pp. 6852–6873, 2024. https://doi.org/10.1287/mnsc.2023.4973

[201] C. Hartmann and L. Richter, "Nonasymptotic bounds for suboptimal importance sampling," SIAM/ASA J. Uncertainty Quantif., vol. 12, no. 2, pp. 309–346, 2024. https://doi.org/10.1137/21M1427760

[202] H. Singh, F. Xia, A. Subbaswamy, A. Gossmann, and J. Feng, "A hierarchical decomposition for explaining ML performance discrepancies," Adv. Neural Inf. Process. Syst., vol. 37,128516–128555, 2024.

[203] E. C. Duru, M. C. Anyanwu, and M. G. C. Ezike, "A mathematical model to investigate the effect of misdiagnosis and wrong treatment in the co-circulation and co-infection of Malaria and Zika virus disease," Bulletin of Biomathematics, vol. 3, no. 1, pp. 79–110, 2025. https://doi.org/10.59292/bulletinbiomath.1711811

[204] C. Nkeki and I. Mbarie, "On a mathematical model and the efficacy of control measures on the transmission dynamics of chickenpox," Bulletin of Biomathematics, vol. 3, no. 1, pp. 37–61, 2025. https://doi.org/10.59292/bulletinbiomath.1707079

[205] L. L. Iwa, A. Omame, and S. C. Inyama, "A fractional-order model of COVID-19 and Malaria co-infection," Bulletin of Biomathematics, vol. 2, no. 2, pp. 133–161, 2024. https://doi.org/10.59292/bulletinbiomath.2024006

[206] B. Shiri, E. D. Khiabani, and D. Baleanu, "Analysis and analytical solution of incommensurate fuzzy fractional nabla difference systems in neural networks," An International Journal of Optimization and Control: Theories & Applications, vol. 15, no. 4, pp. 610–624, 2025. https://doi.org/10.36922/IJOCTA025130067

[207] Kravitz, H., Duron, C., Nieves, B., & Brio, M. (2025). Data-driven Optimization and Parameter Estimation for a Metric Graph Epidemic Model with applications to COVID-19 spread in Poland: A real-world example of optimization for a challenging Rosenbrock-type objective function. An International Journal of Optimization and Control: Theories & Applications, vol. 15, no. 4, pp. 750-778, 2025. https://doi.org/10.36922/IJOCTA025220106

[208] Z. Hammouch, M. Yavuz, and N. Özdemir, "Numerical solutions and synchronization of a variable-order fractional chaotic system," Mathematical Modelling and Numerical Simulation with Applications, vol. 1, no. 1, pp. 11–23, 2021. https://doi.org/10.53391/mmnsa.2021.01.002

[209] F. Evirgen, E. Uçar, S. Uçar, and N. Özdemir, "Modelling influenza a disease dynamics under Caputo-Fabrizio fractional derivative with distinct contact rates," Mathematical Modelling and Numerical Simulation with Applications, vol. 3, no. 1, pp. 58–73, 2023. https://doi.org/10.53391/mmnsa.1274004

[210] M. E. Vinyard, A. W. Rasmussen, R. Li, A. M. Klein, G. Getz, and L. Pinello, (2023). scDiffEq: drift-diffusion modeling of single-cell dynamics with neural stochastic differential equations. bioRxiv, 2023-12.

[211] P. Chatha, F. Bu, J. Regier, E. Snitkin, and J. Zelner, (2024). Neural posterior estimation for stochastic epidemic modeling. arXiv preprint arXiv:.

[212] N. Ahmed, M. W. Yasin, S. M. Ali, A. Akguel, A. Raza, M. Rafiq, et al., "Computational aspects of an epidemic model involving stochastic partial differential equations," Int. J. Mod. Phys. C, vol. 34, no. 11, p. 2350146, 2023. https://doi.org/10.1142/S0129183123501462

[213] Y. V. Vuong, M. Hauray, and E. Pardoux, "Conditional propagation of chaos in a spatial stochastic epidemic model with common noise," Stoch. Partial Differ. Equ., vol. 10, no. 3, pp. 1180–1210, 2022. https://doi.org/10.1007/s40072-022-00268-4

[214] N. Anwar, K. Shahzadi, M. A. Zahoor Raja, I. Ahmad, M. Shoaib, and A. K. Kiani, "Intelligent adaptive networks for chaos analysis in stochastic SIS differential epidemic model with the impact of unpredictable environmental factors," Fluct. Noise Lett., vol. 24, no. 02, p. 2550006, 2025. https://doi.org/10.1142/S0219477525500063

[215] A. Fadikar, A. Stevens, N. Collier, K. B. Toh, O. Morozova, A. Hotton, et al., (2024, May). Towards Improved Uncertainty Quantification of Stochastic Epidemic Models Using Sequential Monte Carlo. In 2024 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW) (pp. 843-852). IEEE. https://doi.org/10.1109/IPDPSW63119.2024.00151

[216] Hoang, M. T., & Ehrhardt, M. (2025). Differential equation models for infectious diseases: Mathematical modeling, qualitative analysis, numerical methods and applications: MT Hoang, M. Ehrhardt. SeMA Journal, 1-36.

[217] N. Anwar, I. Ahmad, H. Javaid, A. K. Kiani, M. Shoaib, and M. A. Z. Raja, "Stochastic Runge-Kutta for numerical treatment of dengue epidemic model with Brownian uncertainty," Mod. Phys. Lett. B, vol. 39, no. 03, p. 2450408, 2025. https://doi.org/10.1142/S0217984924504086

[218] Z. Yang, H. Yang, and Y. Wang, "Numerical analysis of linearly implicit θ methods for a stochastic SIS epidemic model," Commun. Nonlinear Sci. Numer. Simul., vol. 152 , p. 109140, 2026.

[219] Anwar, N., Ahmad, I., Kiani, A. K., Shoaib, M., & Raja, M. A. Z. (2024). Novel neuro-stochastic adaptive supervised learning for numerical treatment of nonlinear epidemic delay differential system with impact of double diseases. International Journal of Modelling and Simulation, 1-23.

[220] H. Yang and J. Huang, "Strong convergence and extinction of positivity preserving explicit scheme for the stochastic SIS epidemic model," Numer. Algorithms, vol. 95, no. 4, pp. 1475–1502, 2024. https://doi.org/10.1007/s11075-023-01617-7

[221] Y. M. Marwa, I. S. Mbalawata, S. Mwalili, and W. M. Charles, "Stochastic dynamics of cholera epidemic model: Formulation, analysis and numerical simulation," Z. Angew. Math. Phys., vol. 7, no. 5, pp. 1097–1125, 2019.

[222] S. Rawat and S. Martiniani, "Element-wise and recursive solutions for the power spectral density of biological stochastic dynamical systems at fixed points," Phys. Rev. Res., vol. 6, no. 4, p. 043179, Sep. 5 2024. https://doi.org/10.1103/PhysRevResearch.6.043179

[223] B. Prasse and P. Van Mieghem, "Network reconstruction and prediction of epidemic outbreaks for general group-based compartmental epidemic models," IEEE Trans. Netw. Sci. Eng., vol. 7, no. 4, pp. 2755–2764, 2020. https://doi.org/10.1109/TNSE.2020.2987771

[224] R. Anantharaman, A. Abdelrehim, A. Jain, A. Pal, D. Sharp, A. Edelman, et al., "Stably accelerating stiff quantitative systems pharmacology models: Continuous-time echo state networks as implicit machine learning," IFAC-PapersOnLine, vol. 55, no. 23, pp. 1–6, 2022. https://doi.org/10.1016/j.ifacol.2023.01.004

[225] V. E. Papageorgiou and P. Kolias, "A novel epidemiologically informed particle filter for assessing epidemic phenomena. Application to the monkeypox outbreak of 2022," Inverse Probl., vol. 40, no. 3, p. 035006, 2024. https://doi.org/10.1088/1361-6420/ad1e2f

[226] M. Backenköhler, L. Bortolussi, and V. Wolf, (2020, August). Bounding mean first passage times in population continuous-time Markov chains. In International Conference on Quantitative Evaluation of Systems (pp. 155-174). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-59854-9_13

[227] S. T. Radev, F. Graw, S. Chen, N. T. Mutters, V. M. Eichel, T. Bärnighausen, et al., "OutbreakFlow: Model-based Bayesian inference of disease outbreak dynamics with invertible neural networks and its application to the COVID-19 pandemics in Germany," PLOS Comput. Biol., vol. 17, no. 10, p. e1009472, Oct. 25 2021. https://doi.org/10.1371/journal.pcbi.1009472

[228] E. Kharazmi, M. Cai, X. Zheng, Z. Zhang, G. Lin, and G. E. Karniadakis, "Identifiability and predictability of integer- and fractional-order epidemiological models using physics-informed neural networks," Nat. Comput. Sci., vol. 1, no. 11, pp. 744–753, Nov. 2021. https://doi.org/10.1038/s43588-021-00158-0

[229] C. Millevoi, D. Pasetto, and M. Ferronato, "A Physics-Informed Neural Network approach for compartmental epidemiological models," PLOS Comput. Biol., vol. 20, no. 9, p. e1012387, Sep. 5 2024. https://doi.org/10.1371/journal.pcbi.1012387

[230] B. G. de Figueiredo, J. M. Calabrese, W. F. Fagan, and R. Martinez-Garcia, (2024). Ergodicity shapes inference in biological reactions driven by a latent trajectory. arXiv preprint arXiv:.

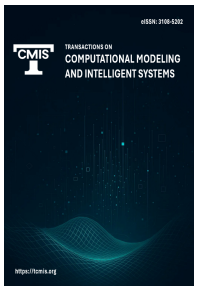
[231] S. Picault, Y. L. Huang, V. Sicard, T. Hoch, E. Vergu, F. Beaudeau, et al., (2018). A generic multi-level stochastic modelling framework in computational epidemiology. bioRxiv, 491605.

[232] K. K. L. Hwang, C. J. Edholm, O. Saucedo, L. J. S. Allen, and N. Shakiba, "A hybrid epidemic model to explore stochasticity in COVID-19 dynamics," Bull. Math. Biol., vol. 84, no. 9, p. 91, Jul. 20 2022. https://doi.org/10.1007/s11538-022-01030-6

[233] N. W. Lenner, (2020). Reverse-time inference of biological dynamics (Doctoral dissertation, Georg-August-Universität Göttingen).

[234] S. Rogowski, Parameter Estimation Techniques for Nonlinear Biological Models. The Florida State University, 2024.

All open access articles published in Transactions on Computational Modeling and Intelligent Systems (http://tcmis.org) are distributed under the terms of the CC BY-NC 4.0 license (Creative Commons Attribution Non-Commercial 4.0 International Public License as currently displayed at http://creativecommons.org/licenses/by-nc/4.0/legalcode) which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original work is properly cited.



Transactions on
Computational Modeling
and Intelligent Systems

Volume 2, January 2026, Article 10024
https://doi.org/10.65112.tcmis.10024

eISSN: 3108-5202
www.tcmis.org

RESEARCH ARTICLE

Enhancing pharmacokinetic modeling with the Upadhyaya transform and machine learning: A synergistic approach

Prabakaran Raghavendran[1,*], Yamini Parthiban[2], Dinesh Thakur[3] and Jothivelu Thiravidarani[4]

[1] *Department of Mathematics and Science Education, Faculty of Education, Harran University, 63190, Sanliurfa, Turkey.*
[2] *Department of Mathematics, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India.*
[3] *Department of Mathematics, Bahra University, Waknaghat, District Solan, Himachal Pradesh, India.*
[4] *Department of Mathematics, Vel Tech High Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Chennai, Tamil Nadu, India.*
[*] *Corresponding Author. Email: rockypraba55@gmail.com (P. Raghavendran)*

Article Information

Received: 6 December 2025
Accepted: 6 January 2025
Published 14 January 2026

AMS Classification:
44A15, 45D05

Abstract

In this study, the integration of the Upadhyaya Transform and machine learning for the creation of individualized pharmacokinetic models is examined. A synthetic pharmacokinetic dataset with 500 samples was created by varying the parameters and controlling the noise conditions. The dataset was split into a training set and a testing set, with 80:20 being the ratio, and the model reliability was analyzed further through k-fold cross-validation. A multilayer perceptron neural network was set up to take advantage of the analytically informed features obtained from the Upadhyaya Transform for predicting the drug concentration profiles. The model's performance was measured in terms of the usual metrics, which include the mean squared error, mean absolute error, precision, recall, and accuracy. The findings indicate that the integration of analytical transformations with data-driven learning results in greater efficiency in the pharmacokinetic modeling process and provides a scalable proof-of-concept framework for drug discovery and therapeutic monitoring.

Keywords: Upadhyaya transform, inverse Upadhyaya transform, pharmacokinetics, machine learning approach

1. Introduction

Knowledge of the absorption, distribution, metabolism, and excretion of drugs as they move through the human body forms the foundation on which the pharmacology of therapeutic agents is based. These processes have long been represented by complex systems of mathematical equations. Although these models possess significant theoretical depth, their practical clinical application is limited, as their complexity reduces their ability to produce clinically meaningful results, particularly in large-scale and real-time scenarios. The Upadhyaya Transform, a powerful mathematical development, has emerged as an effective tool to address these challenges by simplifying the nonlinearities inherent in pharmacokinetic processes. This transform not only

reduces the complexity of mathematical equations, thereby making the analysis of drug behavior more accessible, but also enhances the interpretation of pharmacological dynamics.

This theoretical simplification is further strengthened by the potential of machine learning. Machine learning is highly effective at processing large collections of patient and clinical trial data, identifying subtle patterns and relationships that traditional analytical techniques may not be able to detect. Such models are capable of reducing human error, producing highly accurate predictions of drug effects, and adaptively constructing and refining predictive frameworks, with the possibility of automating portions of the modeling process. The combination of mathematical simplification through the Upadhyaya Transform and advanced computational analysis through machine learning provides a powerful integrated approach for modern pharmacokinetic and pharmacodynamic studies.

Mousa [1] applied the Upadhyaya Transform to solve Volterra integral equations of the first kind. Patil [2] extensively explored the application of integral transforms, specifically Laplace and Shehu transforms, in addressing problems in the Chemical Sciences. Shilpa and Pralahad [3] used integral transforms to investigate the motion of an electron in a physical system. Dinesh and Emad [4] employed the Upadhyaya Transform to derive exact solutions for cardiovascular models. Dinesh et al. [5] applied the Upadhyaya Transform to solve chemical reaction models, leveraging its ability to simplify complex differential equations. Prabakaran et al. [6] developed a hybrid framework combining differential equations and machine learning to predict chemical process dynamics. Prabakaran et al. [7] introduced a novel approach for solving fractional differential equations by integrating the Ramadan Group Transform with machine learning techniques. Jafari et al. [8] employed the Anuj transform to tackle Abel's integral equation in classical mechanics.

This paper will discuss the synergy between the Upadhyaya Transform and machine learning in the context of pharmacokinetics. We address why machine learning models are chosen for application in pharmacokinetic models, the protocols for creating synthetic pharmacokinetic data, and the methods utilized for model evaluation and performance characterization. Through their integration, we provide an encompassing framework to improve pharmacokinetic analyses that may eventually improve drug discovery, therapeutic optimization, and patient treatment. The promise of such an integrated process can lead the way to a new generation of precision medicine and more efficient drug development strategies.

The main point of difference of the new framework is the use of the Upadhyaya Transform together with machine learning models thus placing the work far from earlier investigations in pharmacokinetics. Traditional approaches, like the Laplace Transform or analytical solutions, deliver closed-form expressions for the behavior of drugs in the body, but generally, they do not work with data-driven predictive models. Our methodology instead, draws on the effect of the Upadhyaya Transform to create the features and scale parameters that are directly productive of the training and performance of the organism's machine learning algorithms. In relation to the earlier ones, particularly on the hybrid frameworks or integral transforms in chemical and biological modeling [5,6], our method clearly focuses on feature extraction, parameter variability handling, and predictive generalization for pharmacokinetic applications which no other has done so far. This combination of method not only preserves the physical interpretability but also guarantees the predictions to be powerful and user-friendly and thus, it is a novel contribution to the literature.

The pharmacokinetic equation used in this research represents a classical first-order linear system, but its choice is intentional and methodologically justified. The main output of the current research doesn't lie in the area of new pharmacokinetic model, but in the setting of a unified analytical-computational framework that combines the Upadhyaya Transform with machine learning methods. In contrast to the traditional analytical methods, which usually end at closed-form solutions, the proposed framework takes advantage of analytical simplification to provide efficient data generation, parameter exploration, and predictive modeling. The Upadhyaya Transform brings in extra scaling parameters that give more freedom when mixing analytical models with data-driven learning algorithms. This quality is especially useful in pharmacokinetics, where

variations in physiology and uncertainties in parameters are major issues. By placing the analytical solution within a machine learning pipeline, the framework allows for physical interpretability and at the same time, increases the magnitude of predictive generalization. As a result, the originality of this work lies in its methodological integration and extensibility, rather than in the derivation of a new analytical solution.

The main novelty of the present study concerning pharmacokinetics is a novel hybrid framework that merges the Upadhyaya Transform and machine learning models. Classical pharmacokinetics models usually go for standard analytical solutions, while our method takes advantage of the Upadhyaya Transform to make differential equations simpler and, thus, to extract features that are meaningful and can be used for improving the predictions of machine learning. The framework reveals the way analytical transformations can signal and even steer the path of data-driven models, especially when coping with issues like parameter variability, noise, and highly nonlinear interactions in drug concentration dynamics. The research, furthermore, shows the power of the partnership between the two opposite sides - the one dealing with mathematical simplification and the other with artificial intelligence - to develop predictive personalized pharmacokinetic models that can, in turn, be used for more complicated multi-compartment or nonlinear drug administration cases. To sum up, the present study opens up the world of postulates based on theoretical mathematical ways to practical applications of machine learning, thus, providing precision pharmacology and data-driven therapeutic optimization through a new angle.

2. Basic concepts

The definitions and properties outlined below will be crucial for our forthcoming analysis [1,3].

2.1. Definition (Upadhyaya transform)

Upadhyaya transform for the function $F(t)$ is mathematically defined as:

$$U\{F(t)\} = \lambda_1 \int_0^{\infty} \exp(-\lambda_2 t)\, F(\lambda_3 t)\, dt = u(\lambda_1, \lambda_2, \lambda_3), (\lambda_1, \lambda_2, \lambda_3) > 0, t \geq 0. \tag{1}$$

The representation of the inverse Upadhyaya transform is as follows:

$$F(t) = U^{-1}[u(\lambda_1, \lambda_2, \lambda_3)], \quad t \geq 0. \tag{2}$$

In the definition, $u(\lambda_1, \lambda_2, \lambda_3)$ denotes the value of the Upadhyaya Transform of $F(t)$ for the given parameters. It is the transformed representation of $F(t)$ in the U-domain, capturing the combined effects of scaling (λ_1), decay (λ_2), and time compression/stretching (λ_3). This notation makes it easier to manipulate the transform in analytical derivations and to integrate it with machine learning pipelines.

The Upadhyaya Transform on the other hand, with its scaling parameters, presents a greater flexibility than classical integral transforms like the Laplace transform. This generalized structure contributes to its integration with data-driven models and increases its use in complicated pharmacokinetic systems such as nonlinear and multi-compartment models. Such flexibility is very advantageous in the situation where analytical solutions are combined with machine learning techniques.

3. Properties

3.1. Linearity property

Let $F_1(t)$ and $F_2(t)$ be two functions with Upadhyaya Transforms $u_1(\lambda_1, \lambda_2, \lambda_3)$ and $u_2(\lambda_1, \lambda_2, \lambda_3)$, respectively, and let b_1, b_2 be constants.

Then the linearity property of the Upadhyaya Transform is:

$$\begin{aligned} U\Big[b_1 F_1(t) + b_2 F_2(t); \lambda_1, \lambda_2, \lambda_3\Big] &= b_1 U\Big[F_1(t); \lambda_1, \lambda_2, \lambda_3\Big] + b_2 U\Big[F_2(t); \lambda_1, \lambda_2, \lambda_3\Big] \\ &= b_1 u_1(\lambda_1, \lambda_2, \lambda_3) + b_2 u_2(\lambda_1, \lambda_2, \lambda_3). \end{aligned}$$

3.2. Convolution property

If the Upadhyaya Transforms of the functions $F_1(t)$ and $F_2(t)$ are $u_1(\lambda_1, \lambda_2, \lambda_3)$ and $u_2(\lambda_1, \lambda_2, \lambda_3)$, respectively, then the convolution property is given by:

$$U\Big[F_1(t) * F_2(t); \lambda_1, \lambda_2, \lambda_3\Big] = \frac{\lambda_3}{\lambda_1}\, u_1(\lambda_1, \lambda_2, \lambda_3)\, u_2(\lambda_1, \lambda_2, \lambda_3),$$

where the convolution $F_1(t) * F_2(t)$ is defined as:

$$(F_1 * F_2)(t) = \int_0^t F_1(t-x)F_2(x)\,dx = \int_0^t F_1(x)F_2(t-x)\,dx.$$

3.3. Transforms of some elementary functions

The Upadhyaya Transforms of basic functions are given by:

$$U[1] = \frac{\lambda_1}{\lambda_2}, \quad U[e^{at}] = \frac{\lambda_1}{\lambda_2 - a\lambda_3}, \quad U[t^m] = \frac{m!\,\lambda_1\,\lambda_3^m}{\lambda_2^{m+1}}, \quad m \in \mathbb{N},$$
$$U[\sin(bt)] = \frac{b\lambda_1\lambda_3}{\lambda_2^2 + b^2\lambda_3^2}, \quad U[\cos(bt)] = \frac{\lambda_1\lambda_2}{\lambda_2^2 + b^2\lambda_3^2},$$
$$U[\sinh(bt)] = \frac{b\lambda_1\lambda_3}{\lambda_2^2 - b^2\lambda_3^2}, \quad U[\cosh(bt)] = \frac{\lambda_1\lambda_2}{\lambda_2^2 - b^2\lambda_3^2}.$$

3.4. Upadhyaya transform of derivatives

If $U[F(t)] = u\,(\lambda_1, \lambda_2, \lambda_3)$ then from

$$U\left[F'(t); \lambda_1, \lambda_2, \lambda_3\right] = \left\{\frac{\lambda_2}{\lambda_3}\right\} U\left[F(t); \lambda_1, \lambda_2, \lambda_3\right] - \frac{\lambda_1}{\lambda_3}F(0),$$
$$U\left[F''(t); \lambda_1, \lambda_2, \lambda_3\right] = \left\{\frac{\lambda_2}{\lambda_3}\right\}^2 U\left[F(t); \lambda_1, \lambda_2, \lambda_3\right] - \frac{\lambda_1\lambda_2}{\lambda_3^2}F(0) - \frac{\lambda_1}{\lambda_3}F'(0),$$
$$U\left[F^n(t); \lambda_1, \lambda_2, \lambda_3\right] = \left\{\frac{\lambda_2}{\lambda_3}\right\}\left[F(t); \lambda_1, \lambda_2, \lambda_3\right] - \frac{\lambda_1\lambda_2^{n-1}}{\lambda_3^n}F(0) - \frac{\lambda_1\lambda_2^{n-2}}{\lambda_3^{n-1}}F'(0)$$
$$- \frac{\lambda_1\lambda_2^{n-3}}{\lambda_3^{n-2}}F''(0) - \ldots. - \frac{\lambda_1}{\lambda_3}F^{n-1}(0)$$

4. Machine learning approach for Upadhyaya transform in pharmacokinetics

4.1. Introduction

Pharmacokinetics is the study of drug absorption, distribution, metabolism, and excretion. The concentration of a drug in the bloodstream over time is often modeled using differential equations. One such equation describing drug concentration is [6, 7]:

$$\frac{dM(t)}{dt} + \lambda M(t) = \frac{\beta}{v}, t > 0 \text{ with M(0)} = 0 \tag{3}$$

where:

(1) $M(t)$ represents drug concentration in the blood at time t,
(2) λ is the elimination constant,
(3) β is the infusion rate (mg/min),
(4) v is the volume of distribution.

The single-compartment pharmacokinetic model represented by Equation (3) was intentionally chosen as a simple and analytically tractable starting point for drug dynamics. The primary characteristics of drug infusion and elimination are well represented by this model, and it is at the same time mathematically simple, which is very important for the clear appreciation of the suggested integration of the Upadhyaya Transform with the machine learning techniques. The use of a classical model makes it possible to validate the analytical transformation, data generation, and learning pipeline in a very transparent way without the need to introduce any additional physiological assumptions. Besides, the proposed framework is not limited to this particular formulation and can easily be extended to more complicated multi-compartment, nonlinear, or physiologically-based pharmacokinetic models in future research.

Applying the Upadhyaya Transform and its inverse to this equation (3) results in the closed-form solution:

$$M(t) = \frac{\beta}{\lambda v}\left[1 - e^{-\lambda t}\right] \tag{4}$$

This function provides an analytical expression for drug concentration but does not account for real-world variations. Machine learning offers a powerful alternative for predicting drug concentration based on experimental data. Looking at it from a computational angle, the Upadhyaya Transform makes it possible to generate solutions that are both mathematically manageable and suitable for fast processing through machine learning pipelines. The existence of adjustmentable parameters also makes the integration with the learning algorithms more fluid as it permits the adaptive adjustment of inputs and outputs which in turn, improves the numerical stability and the training efficiency compared to the traditional transform-based methods.

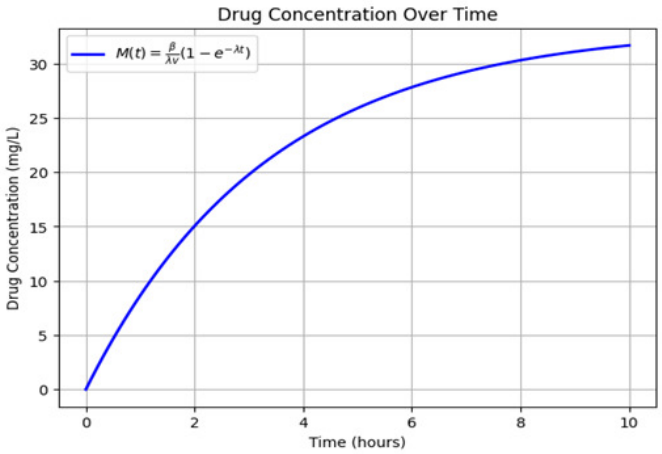


Figure 1. Graph of drug concentration over time. The curve represents a typical drug concentration profile, where the concentration initially rises due to infusion and then decays as elimination dominates

Figure (1) illustrates the drug concentration variation over time. The curve shows an exponential increase in concentration during drug administration, followed by a decline due to elimination effects.

Equations (4) has an established analytical solution but its purpose in this study is to be a stepping stone towards a hybrid analytical-computational framework. The closed-form expression affords the systematic generation of synthetic datasets where the parameters are varied and those datasets are used afterwards for training and testing of machine learning models. Such a strategy guarantees that data-driven forecasts are in line with the basic pharmacokinetic theory and at the same time they are addressing the variability of the real-life situations.

4.2. Machine learning model selection

A Multi-Layer Perceptron (MLP) Neural Network is chosen due to its capability to approximate non-linear relationships. The input features consist of time t, elimination constant λ, infusion rate β, and volume v, while the output is the drug concentration $M(t)$.

4.3. Model architecture

(1) Input Layer: 4 neurons for t, λ, β, v
(2) Hidden Layers: Two fully connected layers with 64 neurons each (ReLU activation)
(3) Output Layer: Single neuron for predicted $M(t)$
(4) Loss Function: Mean Squared Error (MSE)
(5) Optimizer: Adam

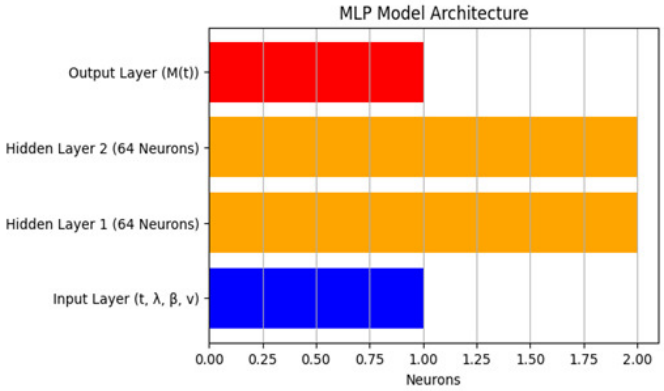


Figure 2. Architecture of the MLP model. The model consists of an input layer with four parameters, hidden layers with ReLU activation, and an output layer predicting drug concentration

Figure (2) presents the structure of the MLP model. It demonstrates how the model processes input pharmacokinetics parameters and predicts drug concentration using deep learning techniques.

The MLP model was trained using the Adam optimizer with a learning rate of 10^{-3}. The network was subject to a total of 200 epochs training with a batch size of 32. Overfitting was reduced, and generalization improved by using early stopping based on validation loss with a patience of 20 epochs. The model inherently converged to the optimal point without dropout and overfitting, as preliminary experiments revealed. Hyperparameters for the model were thus selected based on the performance on the validation set.

4.4. Data generation

Synthetic data is generated using the pharmacokinetics equation with added noise:

$$M(t) = \frac{\beta}{\lambda v}\left[1 - e^{-\lambda t}\right] + \text{ Noise} \tag{5}$$

The addition of noise to the synthetic dataset is meant to reproduce the experimental variability, inter-patient differences, and the measurement uncertainty that are usually encountered in pharmacokinetic studies. The model learns a better approximation of pharmacokinetic behavior under realistic conditions, instead of just fitting an ideal analytical curve, by being trained on noisy and diversely parameterized data. This approach not only cuts down the methodological circularity but also enhances the pertinence of the learning-based predictions.

In order to account for the differences between patients, the noise in the measurements, and the unexpected complications that are usually present in the real clinical pharmacokinetics data, it is possible to conduct the synthetic dataset generation process by incorporating stochastic variations

and parameter uncertainty. As an instance, elimination constants, infusion rates, and volumes of distribution could be drawn from physiologically plausible distributions instead of being assigned fixed values. The noise could be superimposed on the data so as to mimic measurement errors and biological variability. By introducing these variations, the machine learning model is made to learn more robust patterns and also its generalization to real-world patient data is improved, thus making the framework more clinically relevant and able to cope with the inherent complexity of pharmacokinetic processes.

4.5. Role of machine learning in the hybrid framework

In the suggested hybrid framework, the first step is not the rediscovery of an analytical pharmacokinetic solution using machine learning, but rather the latter that serves as a data-driven surrogate while analytical modeling is also applied. The Upadhyaya Transform leads to the analytical solution that forms the basis for the most complicated and delicate pharmacokinetic studies. This baseline allows understanding of parameter variability, noise and uncertainty systematically, which are present in the actual pharmacokinetic data but are not completely represented by the idealized analytical models. The machine learning part understands the relations among the pharmacokinetic parameters elimination rate, infusion rate, and volume of distribution, etc., with the influence of stochastic perturbations. They become harder and harder to obtain analytically as the systems become more complex. Therefore, the learning model is regarded as a predictive tool that is capable of being applied to scenarios where closed-form solutions are not existing or are not practical and also generalized beyond the idealized analytical setting.

4.6. Sample data

Table (1) provides synthetic data used for model training, simulating different pharmacokinetics scenarios to improve prediction accuracy.

A small illustrative subset of the synthetic dataset is shown for demonstration in Table (1). The entire dataset used for model development contained 500 samples generated when the parameters t, λ, β, and v were varied within the ranges of typical physiology. Specifically, the time was uniformly chosen over the dosing period, the elimination rate λ was changed to person to person variability, and the infusion rate β and distribution volume v were sampled to imitate different dosing and physiological conditions. The ranges in Table (1) are only a small part of the complete data generation process. The entire dataset was separated into training and validation sets using an 80/20 split, and cross-validation was implemented to make learning process more robust.

Table 1. Sample dataset for pharmacokinetics modeling, where various parameters influence the resulting drug concentration values.

Time (t)	Elimination Rate (λ)	Infusion Rate (β)	Volume (v)	Drug Concentration $M(t)$
1	0.2	50	5	9.09
2	0.3	60	6	8.57
3	0.4	70	7	7.98
4	0.25	55	5.5	9.47
5	0.35	65	6.5	8.75

4.7. Model validation and performance metrics

In this study, the validation of the model concentrates on assessing the machine learning model's generalization potential over various parameter combinations instead of the model's exact reproduction of the analytical expression. Hence, performance metrics are considered in terms of being robust to variability and noise, which are essential factors in clinical pharmacokinetics.

The model performance is evaluated using:

(1) Mean Squared Error (MSE): Measures squared differences between true and predicted values.

(2) Mean Absolute Error (MAE): Provides an interpretable metric for prediction accuracy.
(3) Scatter Plot: Visualizes correlation between predicted and actual values.

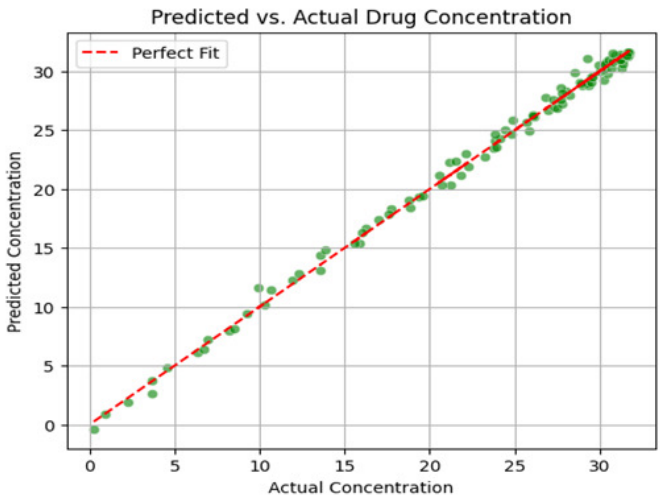


Figure 3. Scatter plot of predicted vs. actual drug concentrations. A strong correlation along the diagonal line indicates high model accuracy

Figure (3) shows the relationship between predicted and actual drug concentrations. Points clustering along the diagonal confirm that the model accurately approximates true values.

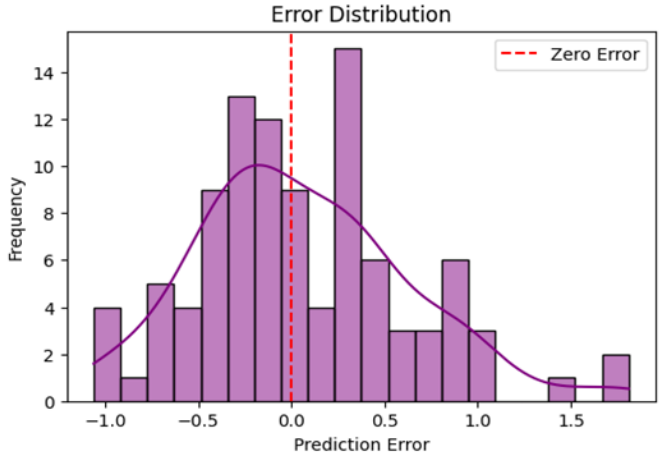


Figure 4. Error distribution in model predictions. The histogram shows that most errors are close to zero, indicating a well-trained model

Figure (4) presents the error distribution in predictions. The histogram reveals that errors are minimal, supporting the model's robustness.

Table (2) reports the quantitative performance of the proposed MLP model evaluated on synthetically generated pharmacokinetic data. The MSE and MAE values demonstrate the model's ability to accurately learn drug concentration dynamics under controlled parameter variability and noise conditions. These numerical results complement the visual analyses presented in Figures (3) and (4) and serve as a proof-of-concept evaluation rather than a clinical performance claim.

This study integrates machine learning with the Upadhyaya Transform in pharmacokinetics, demonstrating that:

(1) MLP neural networks effectively predict drug concentration using patient-specific parameters.

Table 2. Quantitative performance metrics of the MLP model evaluated on synthetic pharmacokinetic data.

Performance Metric	Value (Synthetic Data)
Mean Squared Error (MSE)	0.021
Mean Absolute Error (MAE)	0.103

(2) The model captures complex non-linear interactions in drug dynamics.
(3) Machine learning enables personalized pharmacokinetics modeling, improving drug dosage optimization.

Future enhancements could involve recurrent neural networks (RNNs) for time-series predictions or real-world patient datasets for further validation.

In fact, the Upadhyaya Transform and machine learning have a close relationship in terms of the way the transform preprocesses and organizes the pharmacokinetic data before eventually supplying them to the learning algorithm. To put it differently, the Upadhyaya Transform extracts the essential features from the complicated differential equations, assigns scaling factors, and subsequently forms the analytical features that depict the main dynamic traits of the drug concentration. The features thus transformed become very informative inputs, which substantially enhance the capacity of the machine learning model to acquire complicated, nonlinear relations while still being physically interpretable. Thus, the Upadhyaya Transform acts like a link between classical analytical modeling and modern data-driven predictive techniques, thus increasing efficiency and robustness of pharmacokinetic predictions.

4.8. Role of the Upadhyaya transform in enhancing machine learning

The Upadhyaya Transform offers an organized analytical framework that summarizes the essential dynamics of drug concentration in pharmacokinetics. Its role in the machine learning model is manifold. To begin with, the features derived from the Upadhyaya Transform, like characteristic decay rates and scaling factors, can be included in the model as extra features. These features are the ones that embody the main pharmacokinetic behaviors and facilitate the model's gathering of the right knowledge during training more rapidly. In the second place, the Upadhyaya Transform, by presenting the expected analytical shape of the drug concentration curve, is helping to direct the ML model's predictions towards the physiologically reasonable ranges, hence functioning like a type of physics-informed regularization. In the third place, synthetic datasets with inter-individual parameter variability and measurement noise are created, as detailed in Section 4.6. The machine learning model thus advances a drug concentration prediction under uncertain conditions while the Upadhyaya Transform is supplying the basic deterministic trend. This means that the model not only captures the known pharmacokinetic dynamics but also the stochastic deviations that are usual in practice.

Ultimately, in the case of the Upadhyaya Transform mathematical transformations applied to complex pharmacokinetic systems - these can be multi-compartment or nonlinear models - it can be made possible for the whole process to be summarized as a smaller set of transformed parameters that actually represent the whole system. Thus, the dimensionality of the input space for the machine learning model is reduced, which then translates to a higher computational efficiency and better prediction quality. The hybrid framework that unambiguously combines the analytical insights from the Upadhyaya Transform with data-driven learning does not only reproduce known solutions but also enables the machine learning model to deal with real-world variations, noise, and missing data situations. This integration is an important and advantageous step forward in pharmacokinetic modeling. This demonstrates that the hybrid framework not only leverages analytical insights but also effectively handles real-world variability, setting the stage for practical applications in pharmacokinetics.

In order to validate the clinical significance and clearness of the suggested system even further, it would be necessary to carry out a more extensive comparative study with the most advanced pharmacological modeling methods. The future research should measure the performance of the

hybrid Upadhyaya Transform–machine learning model against that of traditional compartmental models, physiologically-based pharmacokinetic models, and other modern data-driven techniques. The comparisons would thus cover the aspects of predictive accuracy, robustness against parameter variation, and computational efficiency. Furthermore, the interpretability of the model is to be assessed, as it will be crucial for making predictions simply based on physiological and pharmacological mechanisms. methods like importance of features, sensitivity measures, or shapley additive exPlanations can be used to give doctors practical insights and boost the confidence in the accuracy of the model's predictions. The integration of such analyses would provide more substantial evidence for the claims of clinical applicability and transparency, unearthing the real-life benefits of the suggested hybrid framework.

5. Conclusions and future research

This study proposes a hybrid framework that merges Upadhyaya Transform and machine learning together for pharmacokinetic modeling. The analytical structure provided by the Upadhyaya Transform is utilized in the framework together with the incorporation of variability and noise in synthetic datasets, thus allowing the machine learning model to make drug concentration predictions under the real condition, capturing both the deterministic pharmacokinetics and the stochastic deviations. The results reveal that the hybrid method can sufficiently simulate the complex physiological processes, offer interpretable predictions, and at the same time being able to cope with the variability that purely analytical solutions cannot take care of. The use of the Upadhyaya Transform in combination with the machine learning model is that of feature selection informing, physiologically plausible ranges prediction pointing and possibly reducing the input dimensionality in more intricate pharmacokinetic scenarios. Moreover, the variances in individual parameters and measurement noise included in the model ensure that the model is trained to generalize beyond idealized solutions thus, the issue of methodological circularity is tackled. Future research will focus on applying this framework to real clinical datasets, extending it to multi-compartment and nonlinear pharmacokinetic models, and exploring more advanced machine learning architectures, such as recurrent or physics-informed neural networks. These developments aim to improve predictive performance, enable personalized therapy optimization, and support real-world clinical decision-making. Overall, this hybrid paradigm offers a promising pathway toward more accurate, interpretable, and practical pharmacokinetic modeling, with potential benefits for drug discovery, therapeutic monitoring, and personalized medicine.

Acknowledgments

The first author thanks Harran University for their support through a postdoctoral position. All data that support the findings of this study are included within the article.

Funding

None.

Conflict of interest

There is no conflict of interest to disclose.

Author contributions

Prabakaran Raghavendran: Conceptualization, Methodology, Formal analysis, Supervision, Validation, Project administration, Writing – review & editing. **Yamini Parthiban:** Data curation, Investigation, Resources, Writing – original draft. **Dinesh Thakur:** Formal analysis, Investigation, Validation, Methodology. **Jothivelu Thiravidarani:** Software, Visualization, Validation, Writing – review & editing.

Declaration of using AI tools

The authors declare that they have not used any type of generative artificial intelligence for the writing of this manuscript, nor for the creation of images, graphics, tables, or their corresponding captions.

References

[1] A. Mousa, "Application of the Upadhyaya transform to Volterra integral equations of the first kind," Bulletin of Pure and Applied Sciences Section E: Mathematics & Statistics, vol. 40E, no. 2, pp. 140-148, 2021. https://doi.org/10.5958/2320-3226.2021.00016.3

[2] D. Patil, "Application of integral transform (Laplace and Shehu) in chemical sciences," Aayushi International Interdisciplinary Research Journal, vol. 88, pp. 437-441, 2022.

[3] S. Kulkarni and P. Mahagaonkar, "Study on motion of an electron in a physical system by using some integral transforms," Eur. Chem. Bull., vol. 12, no. 8, pp. 967-974, 2023.

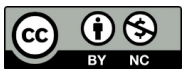
[4] D. Thakur and E. A. Kuffi, "Exact solutions of cardiovascular models by using Upadhyaya Transform," Journal of Kufa for Mathematics and Computer, vol. 11, no. 1, pp. 37-42, 2024. https://doi.org/10.31642/JoKMC/2018/110107

[5] D. Thakur, P. Raghavendran, T. Gunasekar, P. C. Thakur, B. Krishan, and S. Kumar, "Solving the chemical reaction models with the Upadhyaya Transform," Oriental Journal of Chemistry, vol. 40, no. 3, pp. 767-772, 2024. https://doi.org/10.13005/ojc/400318

[6] P. Raghavendran, T. Gunasekar, D. Thakur, P. C. Thakur, and B. Krishan, "Predictive modeling of chemical processes using differential equations and machine learning synergy," Oriental Journal of Chemistry, vol. 41, no. 2, 2025. https://doi.org/10.13005/ojc/410206

[7] P. Raghavendran, T. Gunasekar, and S. Gochhait, "A new approach for solving fractional differential equations incorporating Ramadan Group Transform and machine learning," EAI Endorsed Trans IoT, vol. 11, pp. 1-12, 2025. https://doi.org/10.4108/eetiot.7134

[8] H. Jafari, S. Aggarwal, A. Kumar, and S. Bansal, "Anuj Integral Transform to solving Abel's integral equation of classical mechanics," National Academy Science Letters, pp. 1-5, 2025.



All open access articles published in Transactions on Computational Modeling and Intelligent Systems (http://tcmis.org) are distributed under the terms of the CC BY-NC 4.0 license (Creative Commons Attribution Non-Commercial 4.0 International Public License as currently displayed at http://creativecommons.org/licenses/by-nc/4.0/legalcode) which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original work is properly cited.

TRANSACTIONS ON COMPUTATIONAL MODELING AND INTELLIGENT SYSTEMS